

A Beta Way: A Tutorial For Using Beta Regression in Psychological Research

Jason Geller¹, Robert Kubinec², Chelsea M. Parlett Pelleriti³, and Matti Vuorre⁴

¹Department of Psychology and Neuroscience, Boston College

²University of South Carolina

³Chapman University

⁴Tilburg University

Abstract

Rates, percentages, and proportions are common outcomes in psychology and the social sciences. These outcomes are often analyzed using models that assume normality, but this practice overlooks important features of the data, such as their natural bounds at 0 and 1. As a result, estimates can become distorted. In contrast, treating such outcomes as beta-distributed respects these limits and can yield more accurate estimates. Despite these advantages, the use of beta models in applied research remains limited. Our goal is to provide researchers with practical guidance for adopting beta regression models, illustrated with an example drawn from the psychological literature. We begin by introducing the beta distribution and beta regression, emphasizing key components and assumptions. Next, using data from a learning and memory study, we demonstrate how to fit a beta regression model in R with the Bayesian package `{brms}` and how to interpret results on the response scale. We also discuss model extensions, including zero-inflated, zero- and one-inflated, and ordered beta models. Basic familiarity with regression modeling and R is assumed. To promote wider adoption of these methods, we provide detailed code and materials at <https://doi.org/10.5281/zenodo.1583059>.

Keywords: beta regression, beta distribution, R tutorial, psychology, learning and memory

Jason Geller <https://orcid.org/0000-0002-7459-4505>

Robert Kubinec <https://orcid.org/0000-0001-6655-4119>

Chelsea M. Parlett Pelleriti <https://orcid.org/0000-0001-9301-1398>

Matti Vuorre <https://orcid.org/0000-0001-5052-066X>

The analyses herein were not preregistered. Data, code, and materials for this manuscript can be found at <https://doi.org/10.5281/zenodo.1583059>. The authors have no conflicts of interest to disclose. Author roles were classified using the Contributor Role Taxonomy (CRediT; <https://credit.niso.org/>) as follows: Jason Geller: Conceptualization, Data curation, Formal analysis, Project administration, Resources, Visualization, Writing - original draft; Robert Kubinec: Formal analysis, Validation, Writing - review & editing; Chelsea M. Parlett Pelleriti: Formal analysis, Writing - review & editing; Matti Vuorre: Formal analysis, Writing - review & editing

Correspondence concerning this article should be addressed to Jason Geller, Department of Psychology and Neuroscience, Boston College, McGuinn 300z, Chestnut Hill, MA 2467, USA, Email: drjasongeller@gmail.com

 Preprint

This manuscript is currently **under review** and has not been peer-reviewed. Content is **subject to change**. Please feel free to provide feedback!

1

2

Introduction

3 Many outcomes in psychological research are naturally expressed as proportions or percentages.
 4 These include the proportion of correct responses on a test (e.g., Kornell and Bjork (2008)), the proportion
 5 of time a participant fixates on a stimulus in an eye-tracking task (e.g., James et al., 2025), or the proportion of
 6 respondents endorsing a particular belief (e.g., Costello et al., 2024). A central question for studies involving
 7 proportional data is how to analyze these outcomes in a statistically appropriate and interpretable way.

8 Researchers frequently default to linear models that assume Gaussian (normal) distributions, such
 9 as t-tests, ANOVAs, and linear regression. However, these models make strong assumptions: (1) residuals
 10 are normally distributed, (2) the outcome is unbounded (from $-\infty$ to ∞), and (3) variance is constant across
 11 the range of the data. These assumptions are rarely satisfied in practice (Sladekova & Field, 2024), and they
 12 are especially ill-suited for proportional outcomes, which are bounded between 0 and 1 and often exhibit
 13 heteroscedasticity—non-constant variance, particularly near the boundaries (Ferrari & Cribari-Neto, 2004;
 14 Paolino, 2001; Smithson & Verkuilen, 2006). Violating these assumptions can lead to biased estimates and
 15 spurious inferences, especially when many observations cluster near 0 or 1.

16 In some cases, a generalized linear model (GLM) can relax the assumption of normality. For exam-
 17 ple, binomial and Bernoulli models—often referred to as logistic regression when used with a logit link—are
 18 appropriate when the outcome is binary (0/1) or represents the number of successes out of a fixed number
 19 of trials. These models, however, rely on data that arise from discrete trial structures. They tend to perform
 20 poorly when the observed proportions are truly continuous or when the data show extra variability (overdis-
 21 pension), particularly when many values occur near the boundaries of the scale (0 and 1).

22 The challenges of analyzing proportional data are not new (see Bartlett, 1936). Fortunately, several
 23 existing approaches address the limitations of commonly used models. One such approach is beta regression,
 24 an extension of the generalized linear model that employs the beta distribution (Ferrari & Cribari-Neto, 2004;
 25 Paolino, 2001). Beta regression offers a flexible and robust solution for modeling proportional data directly by
 26 accounting for boundary effects and over-dispersion, making it a valuable alternative to traditional binomial
 27 models. This approach is particularly well-suited for psychological research because it can handle both
 28 the bounded nature of proportional data and the non-constant variance often encountered in these datasets
 29 (Sladekova & Field, 2024). In addition, the direct modeling of proportions allows comparability across tasks
 30 and scales, and can be particularly valuable when only the proportional data is available, as is often the case
 31 with secondary data that lack item-level structure or point values. While usage of these models has faced
 32 obstacles due to theoretical and computational limitations, as we argue in this paper, beta regression and its
 33 extensions now provide an accessible and more robust method to traditional linear modeling.

34 While in this paper we will focus on proportional-responses that lie between 0 and 1—it is important
 35 to note that our analysis applies to any bounded continuous scale. Any bounded scale can be mapped to lie
 36 within 0 and 1 without resulting in a loss of information as the transformation is linear.¹ Consequently, a
 37 scale that has natural end points of -1,234 and +8,451—or any other end points on the real number line short
 38 of infinity—can be modeled using the approaches we describe in this paper.

¹Specifically, for any continuous bounded variable x , we can rescale this variable to lie within 0 and 1 by using the formula $x' = \frac{x - \min(x)}{\max(x) - \min(x)}$ where $0 \leq x' \leq 1$.

39 **A Beta Way Is Possible**

40 With the widespread availability of open-source software such as R (R Core Team, 2024) and its
41 extensive ecosystem of user-developed packages, advanced models like beta regression have become increasingly
42 accessible to applied researchers. Yet, their adoption in psychology remains relatively limited. One
43 contributing factor may be the lack of domain-specific examples that demonstrate how these models address
44 common challenges in psychological data. Although recent years have seen a growing interest in beta regression,
45 and a number of useful tutorials are available (Bendixen & Purzycki, 2023; Coretta & Bürkner, 2025;
46 Heiss, 2021; e.g., Smithson & Verkuilen, 2006; Vuorre, 2019), these resources often either focus on basic
47 implementation or briefly mention extensions without detailing how they can be applied to psychologically
48 relevant research questions.

49 The present tutorial aims to help bridge this gap by providing a comprehensive, applied introduction
50 to beta regression and several of its extensions. In addition to the standard beta model, we walk through zero-
51 inflated, zero-one-inflated, and ordered beta regression. These models are particularly useful for researchers
52 working with proportion outcomes that include boundary values (e.g., exact 0s or 1s) or responses with an
53 inherent ordinal structure. Our goal is to offer practical guidance that enables psychological researchers to
54 implement, interpret, and report these models in ways that directly support their empirical questions.

55 Beyond model specification, we place strong emphasis on interpreting results on the response scale—
56 that is, in terms of probabilities and proportions—rather than relying on often difficult to interpret parameters.
57 This focus makes the models more accessible and meaningful for psychological applications, where effects
58 are often easier to communicate when framed on the original scale of the outcome (e.g., changes in recall
59 accuracy or task performance). Throughout, we provide reproducible code and annotated examples to help
60 readers implement and interpret these models in their own work.

61 We begin the tutorial with a non-technical overview of the beta distribution and its core
62 parameters. We then walk through the process of estimating beta regression models using the R package
63 `{brms}` (Bürkner, 2017), illustrating each step with applied examples. To guide interpretation, we emphasize
64 coefficients, predicted probabilities, and marginal effects calculated using the `{marginaleffects}`
65 package (Arel-Bundock et al., 2024). We also introduce several useful extensions—zero-inflated (ZIB),
66 zero-one-inflated (ZOIB), and ordered beta regression—that enable researchers to model outcomes that in-
67 clude boundary values. Finally, all code and materials used in this tutorial are fully reproducible and
68 available via our GitHub repository: https://github.com/jgeller112/beta_regression_tutorial and on Zenodo
69 (<https://doi.org/10.5281/zenodo.15830595>)².

70 **Beta Distribution**

71 Proportional data pose some challenges for standard modeling approaches: The data are bounded
72 between 0 and 1 and often exhibit non-constant variance (heteroscedasticity) (Ferrari & Cribari-Neto, 2004;
73 Paolino, 2001). Common distributions used within the generalized linear model frameworks often fail to
74 capture these properties adequately, which can necessitate alternative modeling strategies.

²In this article, we try to limit code where possible; however, the online version has all the code needed to reproduce all analyses herein. Furthermore, to promote transparency and reproducibility, the tutorial was written in R version 4.5.1 (R Core Team (2024)) using Quarto (v.1.5.54), an open-source publishing system that allows for dynamic and static documents. This allows figures, tables, and text to be programmatically included directly in the manuscript, ensuring that all results are seamlessly integrated into the document. In addition, we use the `rix` (Rodrigues & Baumann, 2025) R package which harnesses the power of the `nix` (Dolstra & contributors, 2006) ecosystem to help with computational reproducibility. Not only does this give us a snapshot of the packages used to create the current manuscript, but it also takes a snapshot of system dependencies used at run-time. This way reproducers can easily re-use the exact same environment by installing the `nix` package manager and using the included `default.nix` file to set up the right environment. The README file in the GitHub repository contains detailed information on how to set this up to reproduce the contents of the current manuscript, including a video.

75 While we do not have time to delve fully into its derivation, the beta distribution is useful for modeling bounded continuous scales because it is the distribution for the probability of an event. Given that a
 76 probability can take on any value from near 0 (the event will not occur with certainty) to 1 (the event will
 77 occur with certainty), the beta distribution can likewise take on virtually any value in that bounded interval.
 78 As a consequence, the beta distribution is the maximum entropy distribution for *any* bounded continuous
 79 random variable, which means that the beta distribution can represent the full range of possibilities of such a
 80 scale.³ As a consequence, if we have a continuous scale with upper and lower bounds—and no other special
 81 conditions—the beta distribution will in principle provide a very good approximation of the uncertainty of the
 82 scale.

83 Typically, the expected value (or mean) of the response variable is the central estimand scholars want
 84 to estimate. A model should specify how this expected value depends on explanatory variables through two
 85 main components: a linear predictor, which combines the explanatory variables in a linear form ($a + b_1x_1 +$
 86 b_2x_2 , etc.), and a link function, which connects the expected value of the response variable to the linear
 87 predictor (e.g., $E[Y] = g(a + b_1x_1 + b_2x_2)$). In addition, a random component specifies the distribution
 88 of the response variable around its expected value (such as Poisson or binomial distributions, which belong
 89 to the exponential family) (Nelder & Wedderburn, 1972). Together, these components provide a flexible
 90 framework for modeling data with different distributional properties.

91 The beta distribution is continuous and restricted to values between 0 and 1 (exclusive). Its two
 92 parameters—commonly called shape1 (α) and shape2 (β)—govern the distribution’s location, skewness, and
 93 spread. By adjusting these parameters, the distribution can take many functional forms (e.g., it can be sym-
 94 metric, skewed, U-shaped, or even approximately uniform; see Figure 1).

95 To illustrate, consider a test question worth seven points. Suppose a participant scores five out of
 96 seven. The number of points received (5) can be treated as α , and the number of points missed (2) as β . The
 97 resulting beta distribution would be skewed toward higher values, reflecting a high performance (pink line
 98 in Figure 1; “beta(5, 2)”). Reversing these values would produce a distribution skewed toward lower values,
 99 representing poorer performance (orange line in Figure 1; “beta(2, 5)”).

101 I Can’t Believe It’s Not beta

102 While the standard parameterization of the beta distribution uses α and β , a reparameterization to a
 103 mean (μ) and precision (ϕ) is more useful for regression models. The mean represents the expected value
 104 of the distribution, while the dispersion, which is inversely related to variance, reflects how concentrated
 105 the distribution is around the mean, with higher values indicating a narrower distribution and lower values
 106 indicating a wider one. The connections between the beta distribution’s parameters are shown in Equation 1.
 107 Importantly, the variance depends on the average value of the response because uncertainty intervals need to
 108 adjust for how close the value of the response is to the boundary.

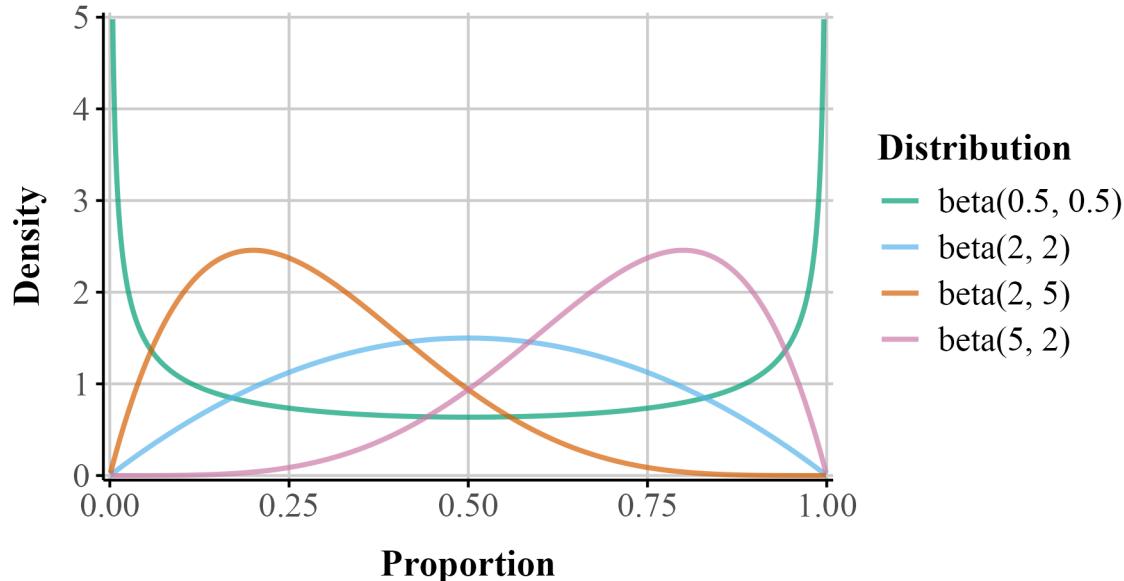
Shape 1: $a = \mu\phi$	Mean: $\mu = \frac{a}{a + b}$	(1)
Shape 2: $b = (1 - \mu)\phi$	Precision: $\phi = a + b$	
	Variance: $var = \frac{\mu \cdot (1 - \mu)}{1 + \phi}$	

109 Thus, beta regression allows modeling both the mean and precision of the outcome distribution. To
 110 ensure that μ stays between 0 and 1, we apply a link function, which allows linear modeling of the mean on
 111 an unbounded scale. A common link-function choice is the logit, but other functions such as the probit or
 112 complementary log-log are possible.

³Technically, this maximum entropy condition is satisfied because the beta(1,1) distribution is uniform over its support. In addition, we assume that the scale has been re-scaled to the [0, 1] interval as we describe above.

Figure 1

beta distributions with different shape1 and shape2 parameters.



113 The logit function, $\text{logit}(\mu) = \log\left(\frac{\mu}{1-\mu}\right)$ links the mean to log-odds which are unbounded, making
 114 linear modeling possible. The logit here no longer carries the same literal *odds* interpretation because there
 115 are no corresponding counts of “successes” and “failures.” Instead, the logit transform here simply maps
 116 the mean of the distribution to the real line. The inverse of the logit, called the logistic function, maps the
 117 linear predictor η back to the original scale of the data $(\mu = \frac{1}{1+e^{-\eta}})$. The coefficients describe how predictors
 118 shift the *average proportion* on the logit scale. Similarly, the strictly positive dispersion parameter is usually
 119 modeled through a log link function, ensuring it remains positive.

120 By accounting for the observations’ natural limits and non-constant variance, the beta distribution
 121 is useful in psychology where outcomes like performance rates or response scales frequently exhibit these
 122 features.

123 Bayesian Approach to Beta Regression

124 Beta regression models can be estimated using either frequentist or Bayesian methods. In this paper,
 125 we adopt a Bayesian framework because it facilitates the estimation and interpretation of more complex
 126 models (Gelman et al., 2013; Johnson et al., 2022; McElreath, 2020). Additionally, the use of Bayesian
 127 statistics in psychology has been steadily growing (Pfadt et al., 2025). In principle, frequentist methods
 128 like maximum likelihood can be framed as Bayesian models with uninformative priors, and as a result, the
 129 modeling perspective we put forward in this paper can apply to either approach. Nonetheless, we note that
 130 in non-linear and hierarchical models, frequentist estimation may require additional adjustments such as
 131 bootstrapping to obtain proper uncertainty intervals, whereas Bayesian modeling handles these extensions
 132 more naturally via exploration of the full joint posterior distribution.⁴

⁴A common concern is that Bayesian methods are slower than frequentist ones. While this is true in general, modern Bayesian computation engines are efficient, and in explanatory modeling our priority is to specify appropriate estimands rather than to optimize for computation speed. Moreover, we use the `{brms}` package (Bürkner, 2017), which provides a high-level interface to Stan (Team, 2023) and uses familiar R formula syntax (similar to `lm()`), making advanced Bayesian modeling accessible to researchers with standard regression backgrounds. The package also supports parallelization, which substantially reduces computation time for large

133 There are several important differences between our Bayesian analysis and the frequentist methods
 134 readers may be more familiar with—most notably, the absence of t - and p -values. To estimate models, the
 135 `{brms}` package uses Stan’s computational algorithms to draw random samples from the posterior distribu-
 136 tion, which represents uncertainty about the model parameters. This posterior is conceptually analogous to
 137 a frequentist sampling distribution. By default, Bayesian models run 4 chains with 2,000 iterations each.⁵
 138 The first 1,000 iterations per chain are warmup and are discarded. The remaining 1,000 iterations per chain
 139 are retained as posterior draws, yielding 4,000 total post-warmup draws across all chains. From these draws,
 140 we can compute the posterior mean (analogous to a frequentist point estimate) and the 95% credible interval
 141 (Cr.I.), which is often compared to a confidence interval.

142 In addition, an important part of Bayesian analyses is prior specification. Priors encode our assump-
 143 tions about plausible parameter values before observing the data and allow the model to regularize estimates,
 144 especially when data are sparse or parameters are weakly identified. To help bridge the conceptual gap for
 145 users more familiar with frequentist models, we begin with the default priors (flat/non-informative, or weakly
 146 informative in some cases) provided by `{brms}`. These priors are intentionally non-informative, and in many
 147 applications produce results that closely align with frequentist estimates, while still offering the flexibility
 148 and interpretive advantages of a Bayesian framework. We strongly urge readers to consider prior specification
 149 strongly in all their work.

150 To ease readers into Bayesian data analysis we provide a metric known as the *probability of direction*
 151 (*pd*), which reflects the probability that a parameter is positive or negative. When a uniform prior is used
 152 (all values equally likely in the prior), *pds* of 95%, 97.5%, 99.5%, and 99.95% corresponds approximately
 153 to two-sided *p*-values of .10, .05, .01, and .001 (i.e., $pd \approx 1 - p/2$ for symmetric posteriors with weak/flat
 154 priors) (see Figure 2 for an illustrative comparison). For directional hypotheses, the *pd* can be interpreted as
 155 roughly equivalent to one minus the *p*-value (Marsman & Wagenmakers, 2016).

156 For reasons of space, we refer readers unfamiliar to Bayesian data analysis to several existing books
 157 on the topic (Gelman et al., 2013; Kruschke, 2015; McElreath, 2020). In addition, we assume readers are
 158 familiar with R, but those in need of a refresher should find Wickham et al. (2023) useful.

159 Beta Regression Tutorial

160 Example Data

161 Throughout this tutorial, we analyze data from a memory experiment examining whether the flu-
 162 ency of an instructor’s delivery affects recall performance (Wilford et al., 2020, Experiment 1A). Instructor
 163 fluency—marked by expressive gestures, dynamic vocal tone, and confident pacing—has been shown to
 164 influence students’ perceptions of learning, often leading learners to rate fluent instructors more favorably
 165 (Carpenter et al., 2013). However, previous research suggests that these impressions do not reliably translate
 166 into improved memory performance (e.g., Carpenter et al., 2013; Toftness et al., 2017; Witherby &
 167 Carpenter, 2022). In contrast, Wilford et al. (2020) found that participants actually recalled more information
 168 after watching a fluent instructor compared to a disfluent one. This surprising finding makes the dataset a
 169 compelling case study for analyzing proportion data, as recall was scored out of 10 possible idea units per
 170 video.

171 In Experiment 1A, ninety-six participants watched two short instructional videos, each delivered
 172 either fluently or disfluently. Fluent videos featured instructors with smooth delivery and natural pacing,
 173 while disfluent videos included hesitations, monotone speech, and awkward pauses. After a distractor task,
 174 participants completed a free recall test, writing down as much content as they could remember from each

datasets.

⁵The Hamiltonian Monte Carlo sampler employed by Stan, which we also use in this paper, can converge with significantly fewer iterations, though rapid convergence depends on model complexity, which is why we use a more conservative standard in this paper.

Figure 2

A Bayesian posterior distribution (assuming a uniform prior) centered at a point estimate chosen so that the probability of direction (pd) equals 0.95, and a frequentist sampling distribution (under the null; centered at 0). In the Bayesian posterior distribution, the blue area represents the pd , and the red area represents the remaining $1 - pd$ of the distribution. In the frequentist sampling distribution, the red tail areas represent the rejection region at $\alpha = 0.10$. In this example, the posterior mean lies exactly at the $1 - \frac{\alpha}{2}$ quantile of the null sampling distribution. For symmetric posteriors with flat priors, the pd is numerically equivalent to the one-sided p -value.

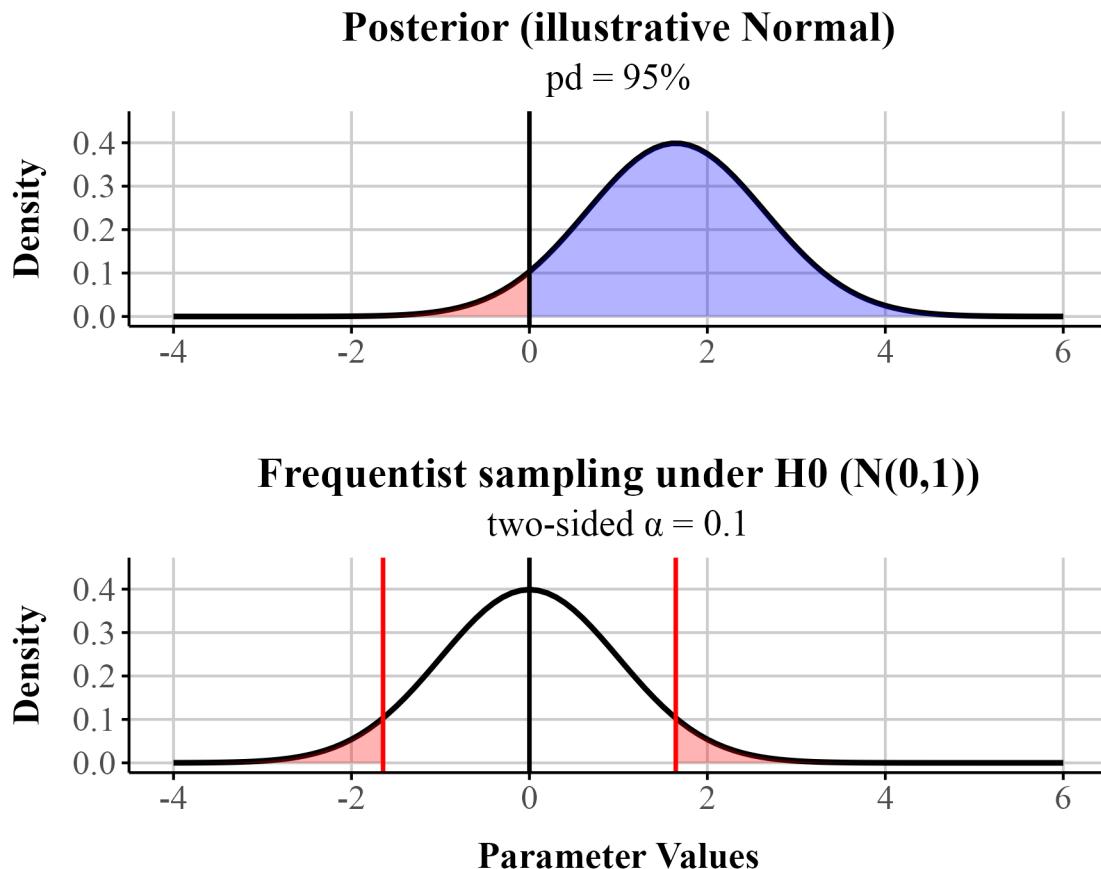


Table 1

Four observations from Wilford et al. (2020). Accuracy refers to the proportion of correctly recalled idea units.

Participant	Fluency	Accuracy
64	Disfluent	0.10
30	Fluent	0.60
12	Fluent	0.10
37	Fluent	0.35

¹⁷⁵ video within a three-minute window. Their recall was then scored for the number of idea units correctly
¹⁷⁶ remembered.

Listing 1 Data needed to run examples

```
# get data here from project folder
fluency_data <- read_csv(here::here("manuscript", "data", "fluency_data.csv"))
```

¹⁷⁷ Our primary outcome variable is the proportion of idea units recalled on the final test, calculated by
¹⁷⁸ dividing the number of correct units by 10. We show a sample of these data in Table 1. The dataset can be
¹⁷⁹ downloaded from GitHub (Listing 1). Because this is a bounded continuous variable (i.e., it ranges from 0 to
¹⁸⁰ 1), it violates the assumptions of typical linear regression models that assume normally distributed residual
¹⁸¹ errors. Despite this, it remains common in psychological research to analyze proportion data using models
¹⁸² that assume normality. In what follows, we reproduce Wilford et al. (2020)'s analysis and then re-analyze
¹⁸³ the data using beta regression and highlight how it can improve our inferences.

184 Reanalysis of Wilford et al. Experiment 1A

¹⁸⁵ In their original analysis of Experiment 1A, Wilford et al. (2020) compared memory performance
¹⁸⁶ between fluent and disfluent instructor conditions using a traditional independent-samples t-test on mean
¹⁸⁷ accuracy for 96 participants. They found that participants who watched the fluent instructor recalled signifi-
¹⁸⁸ cantly more idea units than those who viewed the disfluent version (see Figure 3).

¹⁸⁹ We first replicate this analysis in a regression framework using {brms}. We model final test mean
¹⁹⁰ accuracy—the proportion of correctly recalled idea units across the videos—as the dependent variable. Our
¹⁹¹ predictor is instructor fluency, with two levels: Fluent and Disfluent. We use treatment (dummy) coding,
¹⁹² which is the default in R. This coding scheme sets the first level of a factor (in alphabetical order) as the
¹⁹³ reference level. In this case, Disfluent is the reference, and the coefficient for Fluent reflects the contrast
¹⁹⁴ between fluent and disfluent instructor conditions.

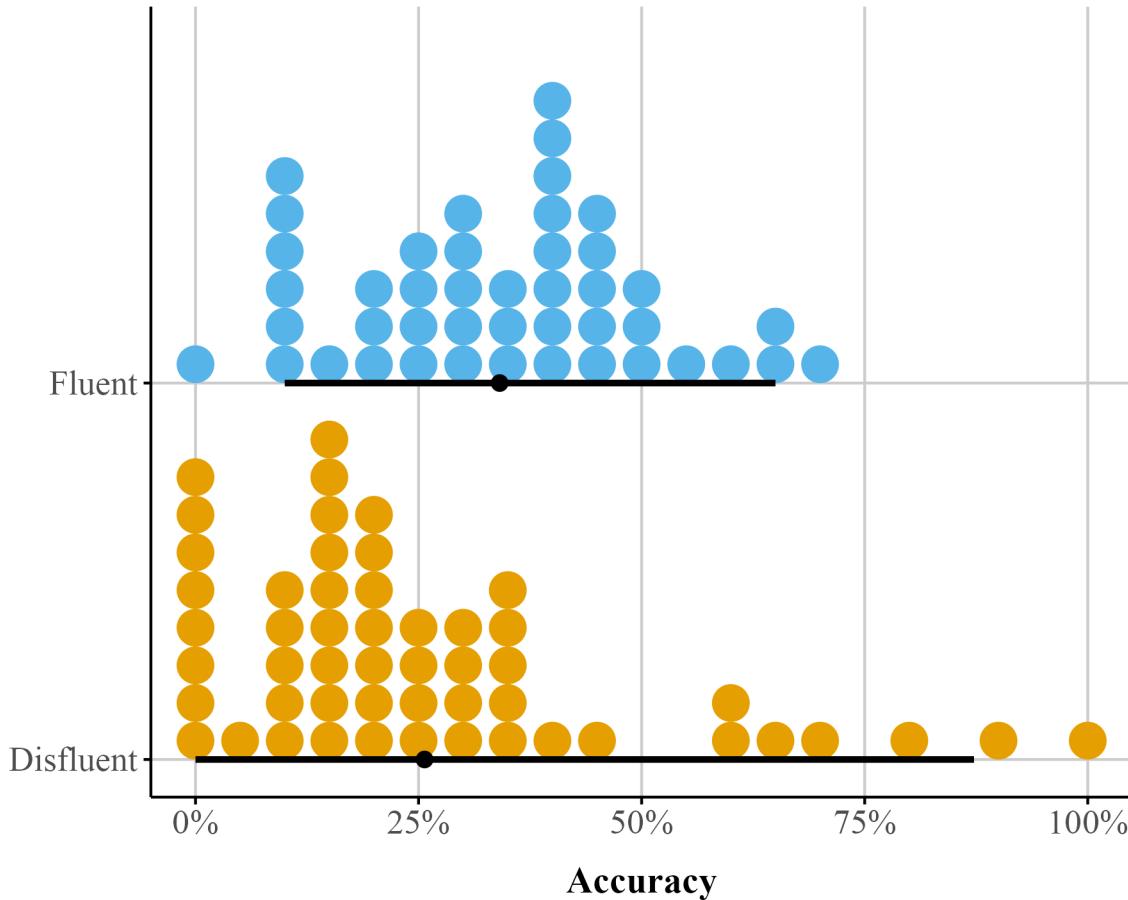
195 Regression Analysis

¹⁹⁶ We first start by loading the {brms} (Bürkner, 2017) and {cmdstanr} (Gabry et al., 2024) packages
¹⁹⁷ (Listing 2). We use the cmdstanr backend for Stan (Team, 2023) because it's faster than the default used to
¹⁹⁸ run models (i.e., rstan),⁶ though all of these models can also be fit with brms defaults.

⁶In order to use the cmdstanr backend you will need to first install the package (see <https://mc-stan.org/cmdstanr/>) and also run cmdstanr::install_cmdstan() if you have not done so already.

Figure 3

Dot plot depicting accuracy distributions for the Fluent and Disfluent conditions. Each condition shows individual data points and summary statistics (mean and 95% percentile interval) to illustrate variability and central tendency.

**Listing 2** Load the {brms} and {cmdstanr} packages

```
library(brms)
library(cmdstanr)
```

199 We fit the model using the `brm()` function from the `{brms}` package (Listing 3). Although not
 200 shown here, we ran the models using four chains (the default), executed in parallel across four cores. When
 201 the model is run in Listing 3, the model summary output will appear in the R console. The output from
 202 `bayes_reg_model` shows each parameter's posterior summary: The posterior distribution's mean and stan-
 203 dard deviation (analogous to the frequentist standard error) and its 95% credible interval, which indicate the
 204 95% of the most credible parameter values. In `{brms}`, the reported Cr.I is an equal-tailed interval, meaning
 205 that the probability mass excluded from the interval is split equally between the lower and upper tails. Ad-
 206 ditionally, the output indicates numerical estimates of the sampling algorithm's performance: Rhat should
 207 be close to one, and the ESS (effective sample size) metrics should be as large as possible given the number
 208 of iterations specified (default is 4000). Generally, ESS ≥ 1000 is recommended (Bürkner, 2017). For the

Listing 3 Fitting a gaussian model with brm().

```

bayes_reg_model <- brm(
  Accuracy ~ Fluency,
  data = fluency_data,
  family = gaussian(),
  file = here::here("manuscript", "models", "model_reg_bayes")
)

```

209 models we present in this paper, convergence is trivial with standard linear models, though we note that these
210 metrics are still important to pay attention to in case of model misfit.

211 Our main question of interest is: what is the causal effect of instructor fluency on final test performance?
212 In order to answer this question, we will have to look at the output summary produced by Listing 3
213 (also see Table 8 under Bayesian LM). the Intercept refers to the posterior mean accuracy in the disfluent
214 condition, $M = 0.257$, as fluency was dummy-coded. The fluency coefficient (FluencyFluent) reflects the
215 mean posterior difference in recall accuracy between the fluent and disfluent conditions: $b = 0.085$. The 95%
216 Cr.I for this estimate spans from -0.003 to 0.169. These values are shown in the “95% Cr.I” columns of the
217 output. These results closely mirror the findings reported by Wilford et al. (2020) (Experiment 1A).

```

218 Family: gaussian
219 Links: mu = identity
220 Formula: Accuracy ~ Fluency
221 Data: fluency_data (Number of observations: 96)
222
223 Regression Coefficients:
224             Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
225 Intercept       0.26      0.03    0.20    0.31 1.00     3993    2515
226 FluencyFluent   0.08      0.04   -0.00    0.17 1.00     4134    2990
227
228 Further Distributional Parameters:
229             Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
230 sigma        0.21      0.02    0.18    0.24 1.00     4035    2970

```

231 The output also includes the effective sample size (ESS) and R (R-hat) values, both of which fall
232 within acceptable ranges, indicating good model convergence. Throughout the tutorial, we focus primarily
233 on posterior mean estimates and their 95% credible intervals. In addition, we report the pd measure in the
234 main summary table (Table 8), provided by the {bayestestR} package (Makowski, Ben-Shachar, Chen, et
235 al., 2019; Makowski, Ben-Shachar, & Lüdecke, 2019). This measure offers an intuitive parallel to p -values,
236 which many readers may find familiar. For example, the fluency effect has a pd of .977, indicating a high
237 probability that the effect is positive rather than negative.

238 Importantly, pd does not indicate whether an effect is meaningfully different from a point value—it
239 only reflects the proportion of the posterior in one direction. To address questions of practical significance,
240 readers can consider the Region of Practical Equivalence (ROPE) with the Cr.Is (Kruschke, 2015). Unlike
241 a hypothesis test of a point null (e.g., exactly zero), the ROPE defines a range of values that are deemed
242 too small to be of substantive importance. As a rule of thumb (see Kruschke, 2018), if more than 95%
243 of the posterior distribution lies inside the ROPE, the effect can be considered practically equivalent to that

²⁴⁴ negligible range. If less than 5% lies inside, the effect can be considered meaningfully different. Intermediate
²⁴⁵ cases are typically labeled undecided.

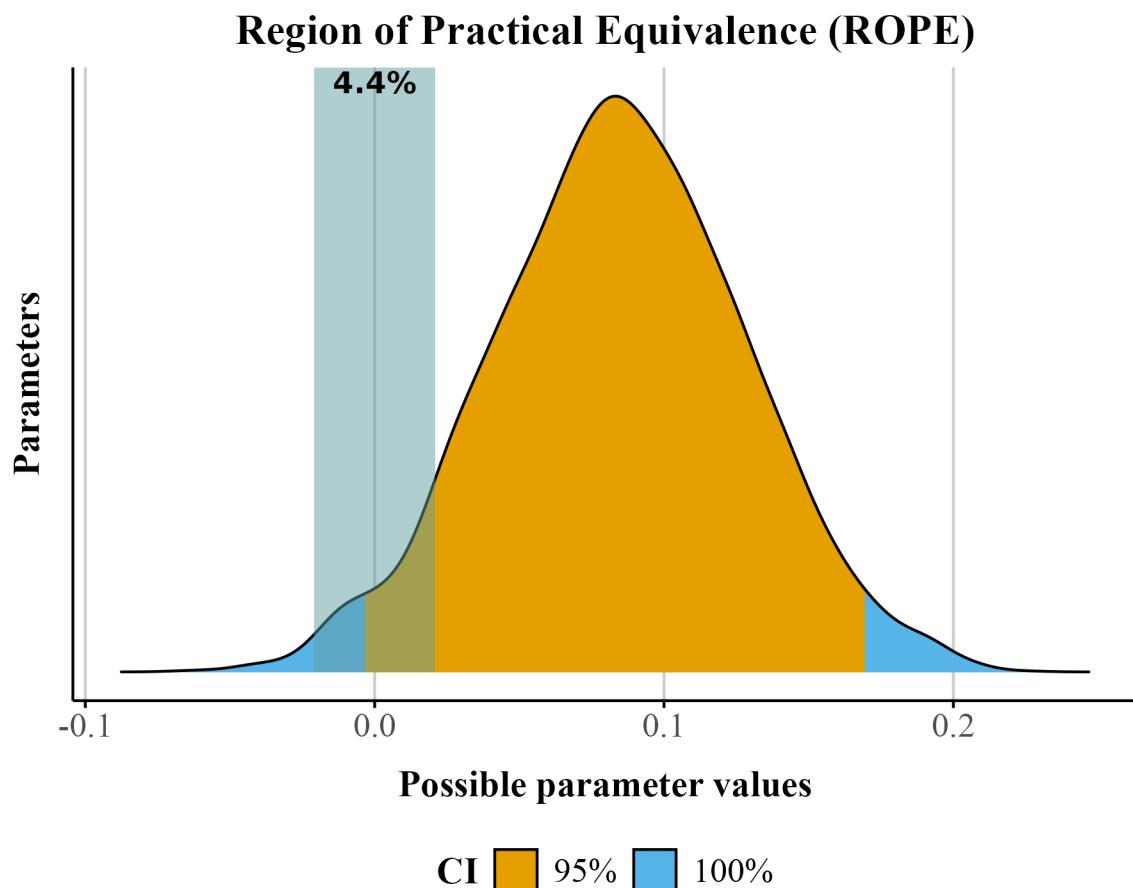
²⁴⁶ The `rope()` function in the `{bayestestR}` package computes the proportion of the posterior within
²⁴⁷ the ROPE to facilitate this evaluation. By default, from bayesian models fit via `{brms}`, the package deter-
²⁴⁸mines a ROPE based on the data (roughly reflecting “negligible” effects), but these defaults should be used
²⁴⁹cautiously. The choice of ROPE should always be guided by theoretical considerations, previous research,
²⁵⁰and the substantive context of the study. In Listing 4, we show how to compute this using `{bayestestR}`. Run-
²⁵¹ning the function with default settings suggests that less than 5% lies within the default ROPE (indicating
²⁵²the effect is larger than .02) (see Figure 4).

Listing 4 Getting ROPE from `bayes_reg_model` object using `rope` function from `{bayestestR}`

```
brms_rope <- bayestestR::rope(bayes_reg_model, ci = .95, ci_method = "ETI")
```

Figure 4

Posterior distribution for the fluency effect showing the ROPE(shaded area) with 95% credible interval (orange) and 100% credible interval (blue). The percentage indicates the proportion of the posterior within the ROPE.



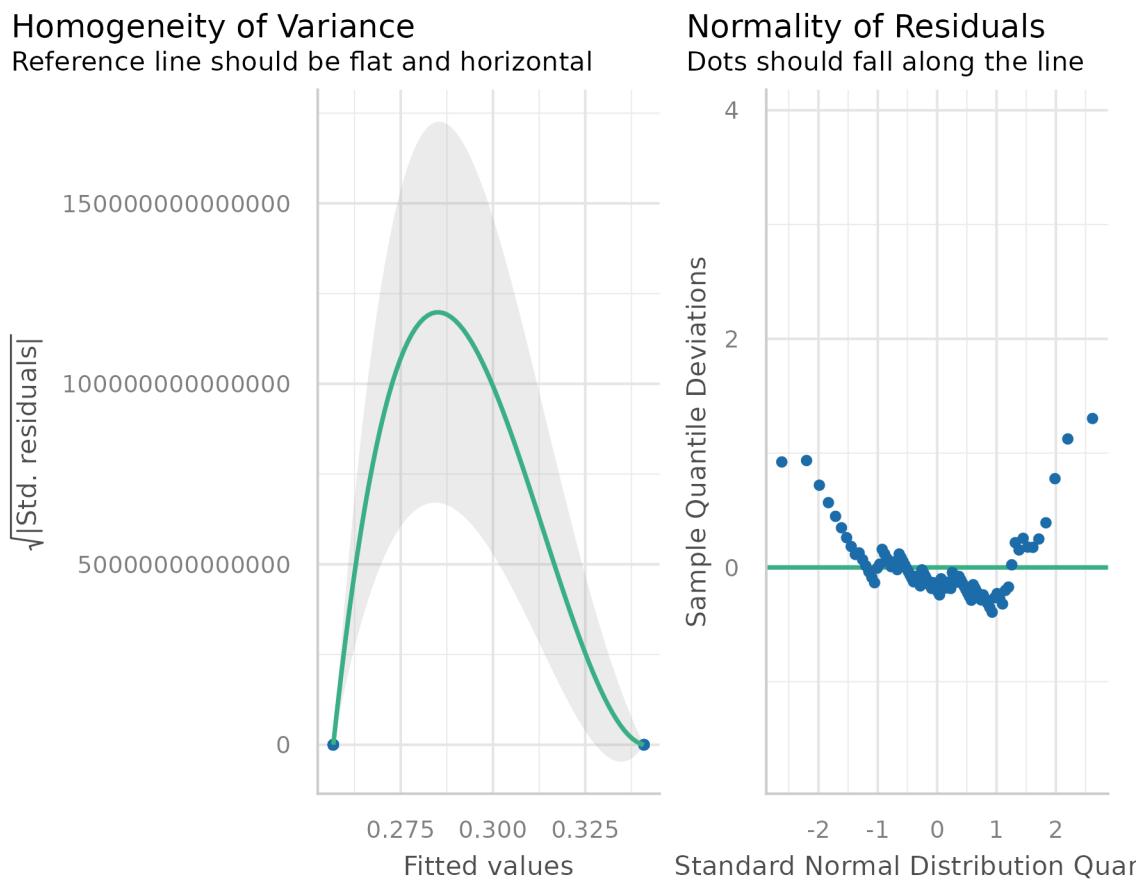
253 **Assumptions.** Wilford et al. (2020) observed that instructor fluency impacts actual learning, us-
 254 ing a standard *t*-test on the mean accuracy. But recall this approach assumes normality of residuals and
 255 homoskedasticity. These assumptions are unrealistic when the response values approach the scale bound-
 256aries (Sladekova & Field, 2024). Does the data we have meet those assumptions? We can use the function
 257 `check_model()` from `{easystats}` (Lüdecke et al., 2022) to check our assumptions easily. The code in List-
 258 ing 5 produces Figure 5. We can see some issues with our model. Specifically, there appears to be violations
 259 of constant variance across the values of the scale (homoskedasticity). In plain terms, this type of model
 260 mis-specification means that a standard OLS model can predict non-sensical values outside the bounds of the
 261 scale.

Listing 5 Checking assumptions with the `check_model()` from `{easystats}` package .

```
check_model(bayes_reg_model, check = c("homogeneity", "qq"))
```

Figure 5

Two assumption checks for our OLS model: Normality (right) and Homoskedasticity (left)



262 We can also examine how well the data fits the model by performing a posterior predictive check
 263 using the `pp_check()` function from `{brms}`. A posterior predictive check involves looking at multiple
 264 model-predicted values and plotting them against the observed data. Ideally, the predicted values (the light

265 blue lines) should show reasonable resemblance with the observed data (dark blue line). In our example
 266 (see Figure 12 (A)) the model-predicted density is slightly too peaked and narrow compared to the data. In
 267 addition, some of the predicted accuracy values are negative.

268 Distributional Regression - Beta

269 It is important to note that there are several justifiable approaches for addressing the distributional
 270 issues observed in the data. For instance, one could analyze median accuracy instead of the mean, use some
 271 type of robust estimator for heterogeneity, or apply non-parametric methods to relax some of the model
 272 assumptions. Alternatively, we can address these issues directly by fitting distributional models (Kneib et
 273 al., 2023; Kruschke, 2013). A key advantage of distributional models is that they are not limited to modeling
 274 only the mean or median of the outcome, but can also model parameters such as the variance (or other shape
 275 parameters) as functions of predictors. This allows examining how instructor fluency may influence not only
 276 average performance, but also the variability in performance across students. If we wanted to keep our mean
 277 accuracy variable and continue to use a Guassian model, we could use a distributional approach and model
 278 the effect of fluency on σ .

279 Given the outcome variable is proportional, another solution would be to run a beta regression model.
 280 Again, we can create the beta regression model in {brms}. In {brms}, we model each parameter indepen-
 281 dently. Recall from the introduction that in a beta model we model two parameters— μ and ϕ . Again we do
 282 this by using the bf() function from {brms} (Listing 6). We specify two formulas, one for μ and one for ϕ
 283 and store it in the model_beta_bayes object below. In the below bf() call, we are modeling accuracy as a
 284 function of fluency only for the μ parameter. For the ϕ parameter, we are only modeling the intercept value.
 285 This is saying dispersion does not change as a function of fluency.

286 To run our beta regression model, we need to exclude 0s and 1s in our data set. If we try to run
 287 a model with our data data_fluency we get an error: `Error: Family 'beta' requires response`
 288 `greater than 0`. This is because the beta distribution only supports observations in the 0 to 1 interval
 289 *excluding exact 0s and 1s*. We need make sure there are no 0s and 1s in our dataset.

290 The dataset contains nine 0s and one 1. One approach is to nudge our 0s towards .01 and our 1s to
 291 .99, or apply a special formula (Smithson & Verkuilen, 2006) so values fall within the [0, 1] interval. We
 292 implore readers not to engage in this practice. Kubinec (2022) showed that this practice can result in serious
 293 distortion of the outcome as the sample size grows larger, resulting in ever smaller values that are “nudged”.
 294 Because the beta distribution is a non-linear model of the outcome, values that are very close to the boundary,
 295 such as 0.00001 or 0.99999, will be highly influential outliers. To run this beta model we will remove the 0s
 296 and 1s, and later in this article we will show how to jointly model these scale end points with the rest of the
 297 data. The model from Listing 6 uses a transformed data_fluency object (called data_beta) where 0s and
 298 1s are removed. When we run this code we should not get an error.

299 **Model Parameters.** In Table 8, under the beta regression column, the coefficient with b_ repre-
 300 sents how fluency of instructor influences the μ parameter estimates (which is the mean of the distribution
 301 here). These coefficients are linear on the logit-scale, but not on the raw accuracy scale. The intercept term
 302 (`b_Intercept`) represents the log odds of the mean on accuracy for the fluent instructor. Log odds that are
 303 negative indicate that it is more likely a “success” (like getting the correct answer) will not happen than that
 304 it will happen. Similarly, regression coefficients in log odds forms that are negative indicate that an increase
 305 in that predictor leads to a decrease in the predicted probability of a “success”.

306 The other component we need to pay attention to is the dispersion or precision parameter coefficients
 307 labeled as phi in Table 8. The dispersion (ϕ) parameter tells us how precise our estimate is. Specifically,
 308 ϕ in beta regression tells us about the variability of the response variable around its mean. Specifically, a
 309 higher dispersion parameter indicates a narrower distribution, reflecting less variability. Conversely, a lower
 310 dispersion parameter suggests a wider distribution, reflecting greater variability. The main difference between

Listing 6 Fitting a beta model without 0s and 1s in brm().

```
# set up model formual
model_beta_bayes <- bf(
  Accuracy ~ Fluency, # fit mu model
  phi ~ 1 # fit phi model
)

# transform 0 to 0.1 and 1 to .99
data_beta <- fluency_data |>
  filter(
    Accuracy != 0,
    Accuracy != 1
  )

beta_brms <- brm(
  model_beta_bayes,
  data = data_beta,
  family = Beta(),
  file = here::here("manuscript","models", "model_beta_bayes_reg_01")
)
```

311 a dispersion parameter and the variance is that the dispersion has a different interpretation depending on the
 312 value of the outcome, as we show below. The best way to understand dispersion is to examine visual changes
 313 in the distribution as the dispersion increases or decreases.

314 Understanding the dispersion parameter helps us gauge the precision of our predictions and the con-
 315 sistency of the response variable. In `beta_brms` we only modeled the dispersion of the intercept. When
 316 ϕ is not specified, the intercept is modeled by default (see Table 8). It represents the overall dispersion in
 317 the outcome across all conditions. Instead, we can model different dispersions across levels of the Fluency
 318 factor. To do so, we add Fluency to the phi model in `bf()`. We model the precision (phi) of the Fluency
 319 factor by using a `~` and adding factors of interest to the right of it (Listing 7).

Listing 7 Fitting beta model with dispersion.

```
model_beta_bayes_disp <- bf(
  Accuracy ~ Fluency, # Model of the mean
  phi ~ Fluency # Model of the precision
)

beta_brms_dis <- brm(
  model_beta_bayes_disp,
  data = data_beta,
  family = Beta(),
  file = here::here("manuscript","models", "model_beta_bayes_dis_run01")
)
```

Table 8 displays the model summary with the precision parameter labeled as phi_Fluency. Since ϕ is modeled on the log scale, the coefficients represent changes in $\log-\phi$ rather than ϕ itself. To see the effect in the original units, we convert the values back by exponentiating. Thus, the effect of the Fluent condition can be understood by comparing the exponentiated predicted ϕ in the Fluent condition to that in the baseline condition.

The ϕ parameters are estimated on the log scale. The term $\beta_{\text{Intercept}}^{(\phi)}$ represents the log-precision for the reference (disfluent) condition. The coefficient $\beta_{\text{FluencyFluent}}^{(\phi)}$ represents the change in log-precision when moving from the disfluent to the fluent condition.

To obtain precision on the original scale, we exponentiate the linear predictor:

$$\phi_{\text{disfluent}} = \exp(\beta_{\text{Intercept}}^{(\phi)}), \quad \phi_{\text{fluent}} = \exp(\beta_{\text{Intercept}}^{(\phi)} + \beta_{\text{Fluency}}^{(\phi)}).$$

The coefficient $\beta_{\text{Fluency}}^{(\phi)}$ therefore describes a *multiplicative* change in precision. Specifically,

$$\frac{\phi_{\text{fluent}}}{\phi_{\text{disfluent}}} = \exp(\beta_{\text{Fluency}}^{(\phi)}).$$

Because the 95% credible interval for $\beta_{\text{Fluency}}^{(\phi)}$ does not include zero, we infer that there is a credible difference in precision between the fluent and disfluent conditions.

It is important to note that these estimates are not the same as the marginal effects we discussed earlier. Changes in dispersion affect the spread or variability of the response distribution without necessarily altering its mean. This makes dispersion particularly relevant for research questions that focus on features of the distribution beyond the average—such as how concentrated responses are. For instance, high dispersion might indicate that individuals cluster at the extremes (e.g., very high or very low ratings), suggesting clustering in the outcome.

A critical assumption of the linear model is homoscedasticity, which means constant variance of the errors. With beta regression model we can include a dispersion parameter for Fluency. Properly accounting for dispersion is crucial because it impacts the precision of our mean estimates and, consequently, the substantive inferences we might make about the coefficients. The inclusion of dispersion in the model increased the uncertainty of the μ coefficient (see Figure 6). This highlights the potential utility of an approach like beta regression over a traditional approach as beta regression can explicitly model dispersion and address issues of heteroscedasticity.

While it is advisable to model precision, if there is uncertainty about the best model, a relatively agnostic approach would be to compare models, for example with leave one out (loo) cross validation, to examine if a dispersion parameter should be considered in our model.⁷

Predicted Probabilities

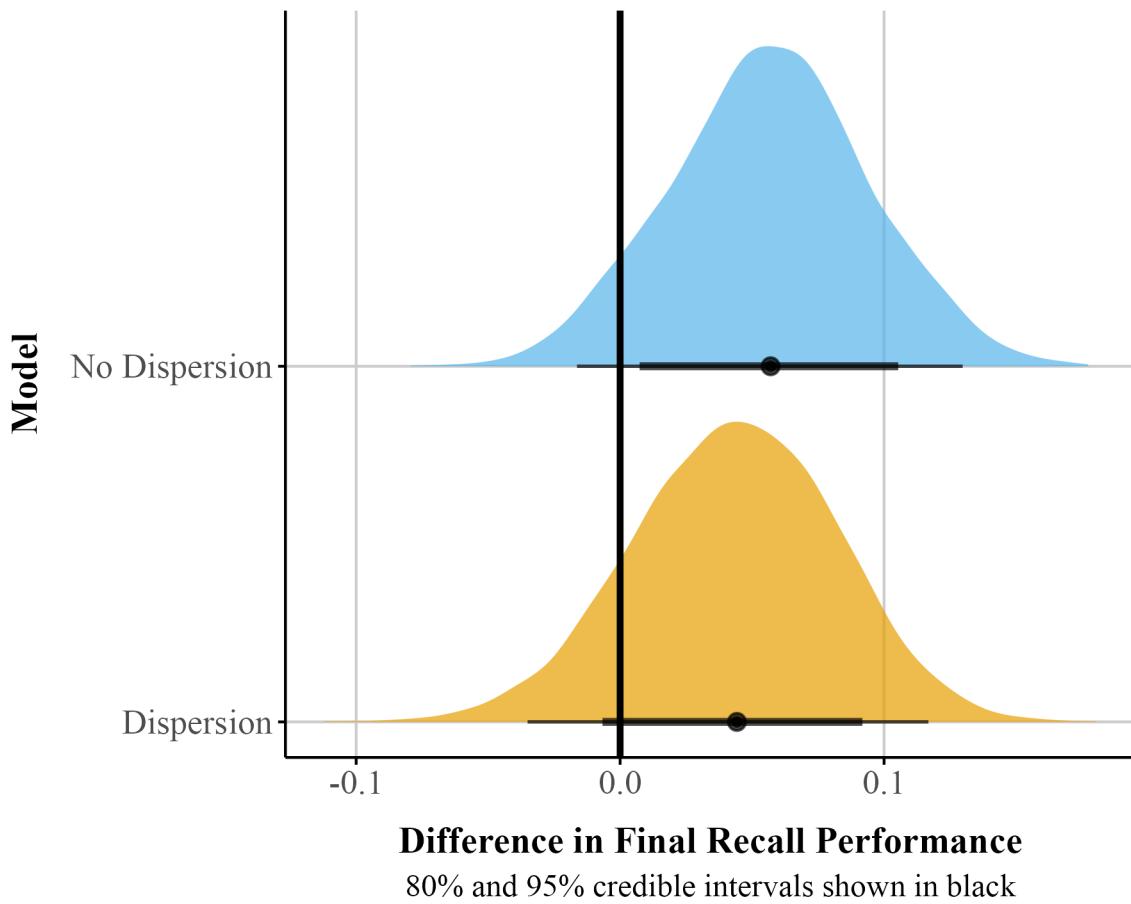
Parameter estimates can be difficult to interpret, and researchers can instead discuss effects on the actual outcome scale (in this case the 0-1 scale). The logit link allows us to transform back and forth between the scale of a linear model and the nonlinear scale of the outcome, which is bounded by 0 and 1. By using the inverse of the logit, we can easily transform our linear coefficients to obtain average effects on the scale of the proportions or percentages, which is usually easier to interpret. In a simple case, we can do this manually, but when there are many factors in your model this can be quite complex.

In our example, we can use the `plogis()` function in base R to convert estimates from the logit scale to the probability scale. The intercept of our model is -0.918, which reflects the logit of the mean accuracy in the disfluent condition. If the estimated difference between the fluent and disfluent conditions is 0.24 on the logit scale, we first add this value to the intercept value (-0.918) to get the logit for the fluent condition:

⁷The model fit statistic LOO-CV can be compared for any set of fitted brms models with the function `loo()`.

Figure 6

Comparison of posterior distributions for the percentage point difference between fluent and disfluent instructors: Simple model (no dispersion for Fluency) vs. complex model with dispersion



359 $-0.83 + 0.20 = -0.63$. We then use `plogis()` to convert both logit values to probabilities (Fluent =
360 35%, Disfluent = 30%).

361 With single coefficients this calculation is trivial, but in more complex models with interactions,
362 it can be quite cumbersome. To help us extract predictions from our model and visualize them we will
363 use a package called `{marginaleffects}` (Arel-Bundock et al., 2024) (see Listing 8). To get the proportions
364 for each of our categorical predictors on the μ parameter we can use the function from the package called
365 `avg_predictions()`. These are displayed in Table 2. These probabilities match what we calculated above.

Listing 8 Load the `{marginaleffects}` package.

```
library(marginaleffects)
options(marginaleffects_posterior_center = mean) # make sure returns mean
```

366 For the Fluency factor, we can interpret the estimates under the Mean column as a proportion or
367 percentage. That is, participants who watched the fluent instructor scored on average 35% on the final exam
368 compared to 30% for those who watched the disfluent instructor. We can also visualize these from `{marginal-`

Listing 9 Predictions from the beta model for each level of Fluency (μ).

```
avg_predictions(
  beta_brms_dis,
  # need to specify the levels of the categorical predictor
  variables = "Fluency"
)
```

Table 2*Predicted probabilities for each level of Fluency (μ).*

Fluency	Mean	95% Cr.I
Disfluent	0.305	[0.252, 0.366]
Fluent	0.348	[0.3, 0.398]

369 effects} using the `plot_predictions()` function (see Listing 10).

Listing 10 Plot predicted probabilities using `plot_predictions()` from {marginaleffects}

```
beta_plot <- plot_predictions(beta_brms_dis, by = "Fluency")
```

370 The `plot_predictions()` function will only display the point estimate with the 95% credible inter-
 371 val. However, Bayesian estimation methods generate distributions for each parameter. This approach allows
 372 visualizing full uncertainty estimates beyond points and intervals. Using the `{marginaleffects}` package, we
 373 can obtain samples from the posterior distribution with the `posterior_draws()` function (see Listing 11).
 374 We can then plot these results to illustrate the range of plausible values for our estimates at different levels
 375 of uncertainty (see Figure 7).

Listing 11 Extracting posterior draws from the beta regression model.

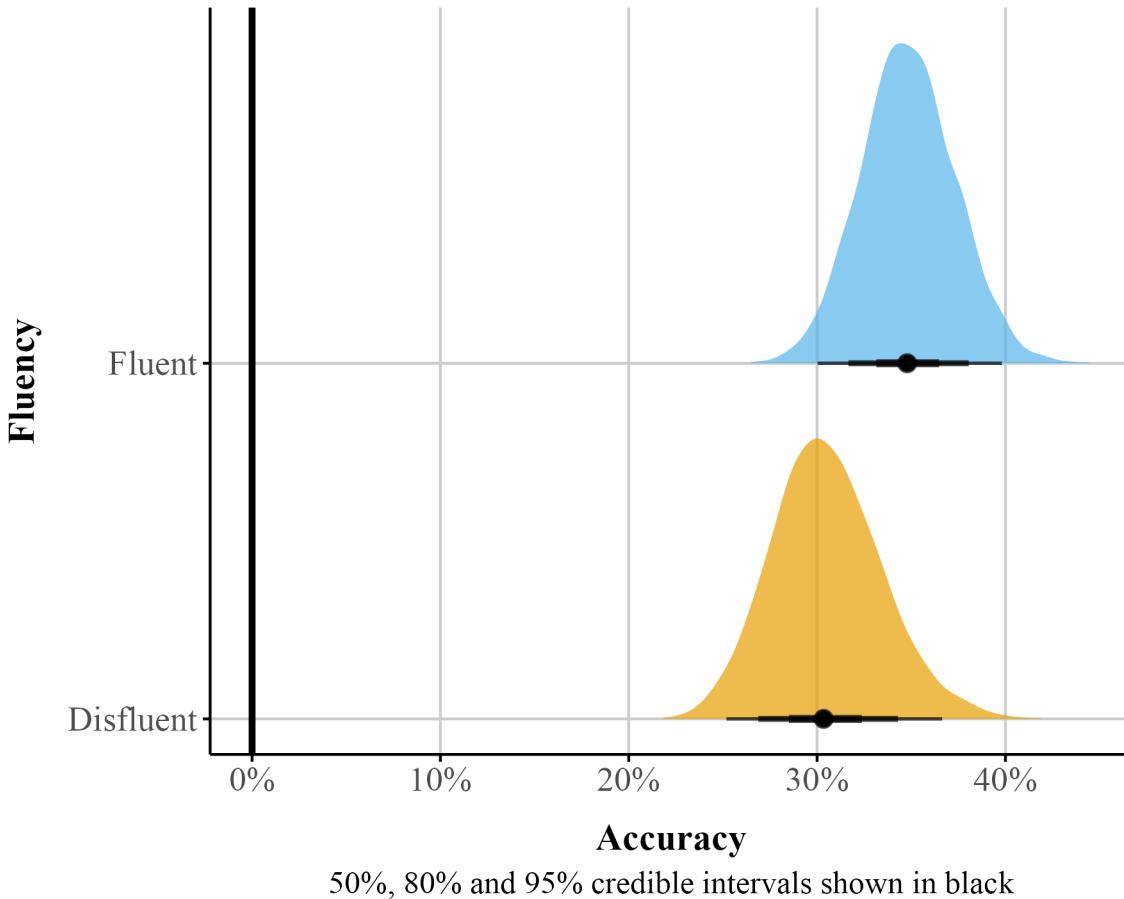
```
# Add a model identifier to each dataset
pred_draws_beta <- avg_predictions(beta_brms_dis, variables = "Fluency") |>
  posterior_draws()
```

376 **Marginal Effects**

377 Marginal effects offer an interpretable way to quantify how changes in a predictor influence an out-
 378 come, while holding other factors constant in a specific manner. In recent years, there has been a thrust to
 379 move away from reporting regression coefficients alone, focusing instead on estimates that are easier to in-
 380 terpret and communicate—particularly in non-linear models (McCabe et al., 2021; Rohrer & Arel-Bundock,
 381 2025). Technically, marginal effects are computed as partial derivatives for continuous variables or as finite
 382 differences for categorical (and sometimes continuous) predictors, depending on the structure of the data and
 383 the research question. Substantively, these procedures translate raw regression coefficients into quantities
 384 that reflect changes in the bounded outcome—for example, an $x\%$ change in the value of a proportion.

Figure 7

Predicted probability posterior distributions by instructor fluency



50%, 80% and 95% credible intervals shown in black

385 There are several types of marginal effects, and their computation can vary across software pack-
 386 ages. For example, the popular `{emmeans}` package (Lenth, 2025) computes marginal effects by holding
 387 all predictors at their means (marginal effects at the mean; MEM). In this tutorial, we use the `{marginalef-`
 388 `fects}` package (Arel-Bundock et al., 2024), which by default computes average marginal effects (AMEs).
 389 AMEs are based on counterfactual predictions: the dataset is conceptually replicated across all unique values
 390 of the predictor of interest, predictions are generated for each row under each counterfactual scenario, and
 391 the resulting differences are then averaged. This approach maintains a strong connection to the observed
 392 data—because predictions are made using each participant’s actual values—while providing a clear and in-
 393 terpretable summary of the effect of interest.

394 One practical use of AMEs is to estimate the average difference between two groups or conditions
 395 which corresponds to the average treatment effect (ATE). Using the `avg_comparisons()` function in the
 396 `{marginaleffects}` package (Listing 12), we can compute this quantity directly. By default, the function returns
 397 the discrete difference between groups. When we take the difference in proportions between two groups it
 398 is called the risk difference. Depending on the audience and modeling goals, the function can also produce
 399 alternative effect size metrics, such as odds ratios or risk ratios. This flexibility makes it a powerful approach
 400 for summarizing and communicating regression results.

401 **μ Difference.** Table 3 presents the estimated difference for the Fluency factor (Mean column).
 402 The difference between the fluent and disfluent conditions is 0.06, indicating that participants who watched

Listing 12 Calculating the difference between probabilities (μ) for instructor fluency with `avg_comparisons()`

```
# get risk difference by default

beta_avg_comp <- avg_comparisons(
  beta_brms_dis,
  variables = "Fluency", # counterfactual difference
  comparison = "difference"
)
```

Table 3

Fluency difference (μ)

Contrast	Mean	95% Cr.I	pd
Fluent - Disfluent	0.043	[-0.035, 0.117]	0.86

403 a fluent instructor scored, on average, 6% higher on the final recall test than those who watched a disfluent
 404 instructor. However, the 95% credible interval includes 0 among the most credible values, suggesting
 405 we cannot rule out the possibility of a null or weakly negative effect. The pd also indicates the effect is
 406 non-significant ($p > .05$). In the basic regression model above, the ROPE value was extracted using the
 407 `{bayestestR}` package.

Listing 13 Extracting ROPE using `{marginaleffects}` for μ difference

```
# get risk difference by default
beta_avg_comp_rope <- avg_comparisons(
  beta_brms_dis,
  variables = "Fluency", # counterfactual difference
  comparison = "difference",
  equivalence=c(-0.02, .02)
)
```

408 The ROPE can also be obtained using the `{marginaleffects}` package by specifying the equivalence
 409 bounds in `avg_comparisons()`. Running the code in Listing 13 returns a ROPE probability of 23%, indicating
 410 that there is insufficient evidence to conclude that the effect exceeds 2% or is practically equivalent.

411 **ϕ difference.** In addition, with `{marginaleffects}`, we can get the actual precision difference between
 412 the two groups on ϕ using similar code to above by setting dpar to “phi” (Listing 14).

413 **Effect Size.** In psychology, it is common to report effect size measures like Cohen’s d (Cohen,
 414 1977). When working with proportions we can calculate something similar called Cohen’s h . Taking our
 415 proportions, we can use the below equation (Equation 2) to calculate Cohen’s h along with the 95% Cr.I
 416 around it. Using this metric we see the effect size is small (0.107), 95% credible interval [-0.002, 0.361].

$$h = 2 \cdot (\arcsin(\sqrt{p_1}) - \arcsin(\sqrt{p_2})) \quad (2)$$

Listing 14 Calculating ϕ difference with avg_comparisons()

```
# get risk difference by default

beta_avg_phi <- avg_comparisons(
  beta_brms_dis,
  variables = "Fluency",
  dpar = "phi",
  comparison = "difference"
)
```

417 **Posterior Predictive Check**

418 Figure 12 (B) shows the predictive check for our beta model. The model's predictions generally
 419 conform to the data as the predictions are now between constrained to the 0-1 interval. However, the model's
 420 predictive performance can be improved if the bounds of the scale are explicitly specified.

421 **Zero-Inflated beta (ZIB) Regression**

422 A limitation of the beta regression model is that it can only accommodate values strictly between
 423 0 and 1—a probability cannot take on values of 0 (the event will not occur with certainty) or 1 (the event
 424 will occur with certainty). In our dataset, we observed 9 rows where Accuracy equals zero. To fit a beta
 425 regression model, we removed these values, but we have left out potentially valuable information from our
 426 model—especially if the end points of the scale are distinctive in some way. In our case, these 0s may be
 427 structural—that is, they represent real, systematic instances where participants failed to answer correctly
 428 (rather than random noise or measurement error). For example, the fluency of the instructor might be a
 429 key factor in predicting these zero responses. We will discuss two approaches for jointly modeling these end
 430 points with the continuous data. First, we can use a zero-inflated beta (ZIB) model. This model still estimates
 431 the mean (μ) and precision (ϕ) of the beta distribution for values between 0 and 1, but it also includes an
 432 additional parameter, α , which captures the probability of observing structural 0s.

433 The zero-inflated beta models a mixture of the data-generating process. The α parameter uses a
 434 logistic regression to model whether the data is 0 or not. Substantively, this could be a useful model when we
 435 think that 0s come from a process that is relatively distinct from the data that is greater than 0. For example,
 436 if we had a dataset with proportion of looks or eye fixations to certain areas on marketing materials, we might
 437 want a separate model for those that do not look at certain areas on the screen because individuals who do
 438 not look might be substantively different than those that look.

439 We can fit a ZIB model using `brms()` and use the `{marginaleffects}` package to make inferences
 440 about our parameters of interest. Before we run a zero-inflated beta model, we will need to transform our
 441 data again and remove the one 1 value in our data—we can keep our 0s. Similar to our beta regression model
 442 we fit in `brms`, we will use the `bf()` function to fit several models. We fit our μ and ϕ parameters as well as
 443 our zero-inflated parameter (α ; here labeled as `zi`). In `brms` we can use the `zero_inflated_beta` family (see
 444 Listing 15).

445 **Posterior Predictive Check**

446 The ZIB model does a bit better at capturing the structure of the data than the beta regression model
 447 (see Figure 12). Specifically, the ZIB model more accurately captures the increased density of values near
 448 the lower end of the scale (i.e., near zero), which the standard beta model underestimates. The ZIB model's

Listing 15 Fitting zib model with `brm()`

```
# keep 0 but remove 1
data_beta_0 <- fluency_data |>
  filter(Accuracy != 1)

# set up model formula for zero-inflated beta in brm
zib_model <- bf(
  Accuracy ~ Fluency, # The mean of the 0-1 values, or mu
  phi ~ Fluency, # The precision of the 0-1 values, or phi
  zi ~ Fluency, # The zero-or-one-inflated part, or alpha
  family = zero_inflated_beta()
)

# fit zib model with brm
fit_zi <- brm(
  formula = zib_model,
  data = data_beta_0,
  file = here::here("manuscript","models", "bayes_zib_model0not1.rds")
)
```

Table 4*Probability fluency difference (μ)*

Contrast	Mean	95% Cr.I	pd
Fluent - Disfluent	0.044	[-0.033, 0.118]	0.875

449 predictive distributions also align more closely with the observed data across the entire range, particularly in
 450 the peak and tail regions. This improved fit likely reflects the ZIB model's ability to explicitly model excess
 451 0s (or near-zero values) via its inflation component, allowing it to better account for features in the data that
 452 a standard beta distribution cannot accommodate.

453 Predicted Probabilities and Marginal Effects

454 Table 8, under the zero-inflated beta regression column, provides a summary of the posterior distribution
 455 for each parameter. As stated before, it is preferable to back-transform our estimates to get probabilities.
 456 To get the predicted probabilities we can again use the `avg_predictions()` and `avg_comparisons()`
 457 functions from `{marginaleffects}` package (Arel-Bundock, 2024) to get predicted probabilities and the probability
 458 difference between the levels of each factor. We can model the parameters separately using the `dpar`
 459 argument setting to: μ , ϕ , α . Here we look at the risk difference for Fluency under each parameter. If one
 460 were interested in the average effect for the entire model, the `dpar` argument could be removed.

461 **Mu.** As shown in Table 4, there is little evidence for an effect of Fluency – the 95% Cr.I includes
 462 zero, suggesting substantial uncertainty about the direction and magnitude of the effect—that is, though most
 463 of the posterior density supports positive effects, nil and weakly negative effects cannot be ruled out.

464 **Dispersion.** As shown in Table 5, the posterior estimates suggest a credible effect of Fluency on dispersion (ϕ), with disfluent responses showing greater variability. The 95% Cr.I for the fluency contrast
 465 does not include zero, indicating a high probability in differences in precision.

Table 5

Probability fluency difference (ϕ)

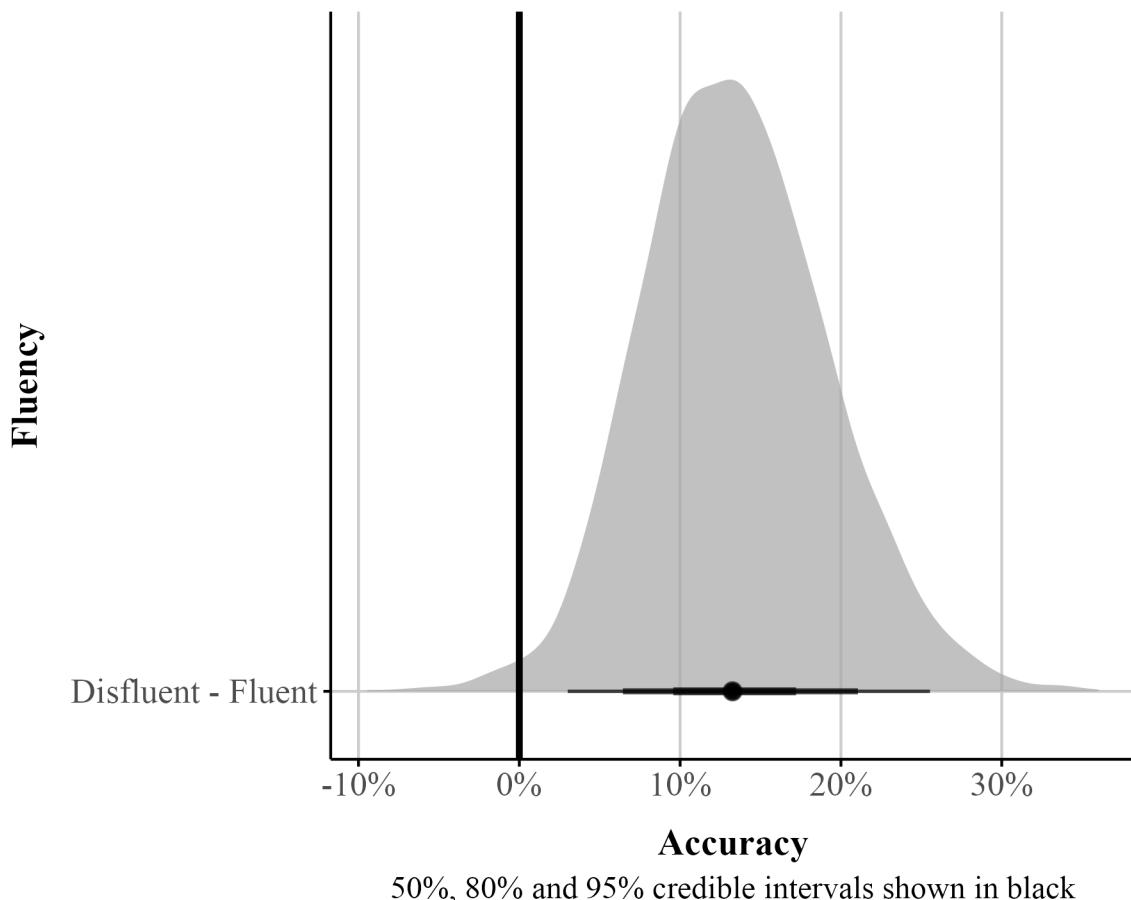
Contrast	Mean	95% Cr.I	pd
Fluent - Disfluent	2.7	[-0.967, 6.664]	0.922

⁴⁶⁷ **Zero-Inflation**

⁴⁶⁸ We can use {marginaleffects} to estimate and plot the posterior difference between the fluent and
⁴⁶⁹ disfluent conditions (see Figure 8). In Figure 8, the posterior distribution for this contrast lies mostly below
⁴⁷⁰ zero, indicating that a fluent instructor is associated with a lower probability of zero responses. The estimated
⁴⁷¹ reduction is approximately 13%. The 95% credible interval does not include zero, which indicates that the
⁴⁷² data provide consistent evidence for a reduction in zero responses under fluent instruction.

Figure 8

Visualization of the predicted difference for zero-inflated part of model



473 **Zero-One-Inflated beta (ZOIB)**

474 The ZIB model works well if there are 0s in your data, but not 1s.⁸ In our previous examples we
 475 either got rid of both 0s and 1s (beta regression), or removed the 1s (ZIB). Sometimes it is theoretically useful
 476 to model both 0s and 1s as separate processes or to consider these values as essentially similar parts of the
 477 continuous response, as we show later in the ordered beta regression model. For example, this is important
 478 in visual analog scale data where there might be a prevalence of responses at the bounds (Kong & Edwards,
 479 2016), in JOL tasks (Wilford et al., 2020), or in a free-list task where individuals provide open responses to
 480 some question or topic which are then recoded to fall between 0-1 (Bendixen & Purzycki, 2023). Here 0s
 481 and 1s are meaningful; 0 means item was not listed and 1 means the item was listed first.

482 Similar to the beta and zero-inflated models discussed above, we can fit a zero-and-one-inflated beta
 483 (ZOIB) model in `{brms}` using the `zero_one_inflated_beta` family. This formulation simultaneously
 484 estimates the mean μ and precision ϕ of the Beta component, as well as two inflation parameters: α , the
 485 probability that an observation is at either boundary (0 or 1), and γ , the conditional probability that, given
 486 an observation falls on a boundary, it takes the value 1 rather than 0. In other words, α determines how often
 487 responses occur exactly at the endpoints, and γ determines the balance between zeros and ones among those
 488 endpoint values. This specification allows the model to capture both the continuous variation in the interior
 489 of the (0, 1) interval and the presence of exact boundary values.

490 To illustrate how α and γ shape the distribution, Figure 9 displays simulated data across a range
 491 of parameter combinations. As α increases, more responses occur at the endpoints. As γ increases, the
 492 proportion of those endpoint responses that are 1 increases relative to 0, producing increasingly pronounced
 493 spikes at 1 as γ approaches 1. Together, these parameters give the ZOIB model the flexibility to represent
 494 datasets with mixtures of continuous values and exact zeros and ones.

495 To fit a ZOIB model we use the `bf()` function. We model each parameter as a function of Fluency.
 496 We then pass the `zoib_model` to our `brm()` function (see Listing 16). The summary of the output is in
 497 Table 8 (under ZOIB).

498 **Model Parameters**

499 The output for the model is lengthy because we are estimating four distinct components, each with
 500 their own independent responses and sub-models. All the coefficients are on the logit scale, except ϕ , which is
 501 on the log scale. Thankfully drawing inferences for all these different parameters, plotting their distributions,
 502 and estimating their average marginal effects looks exactly the same—all the `brms` and `{marginaleffects}`
 503 functions we used work the same.

504 **Predictions and Marginal Effects**

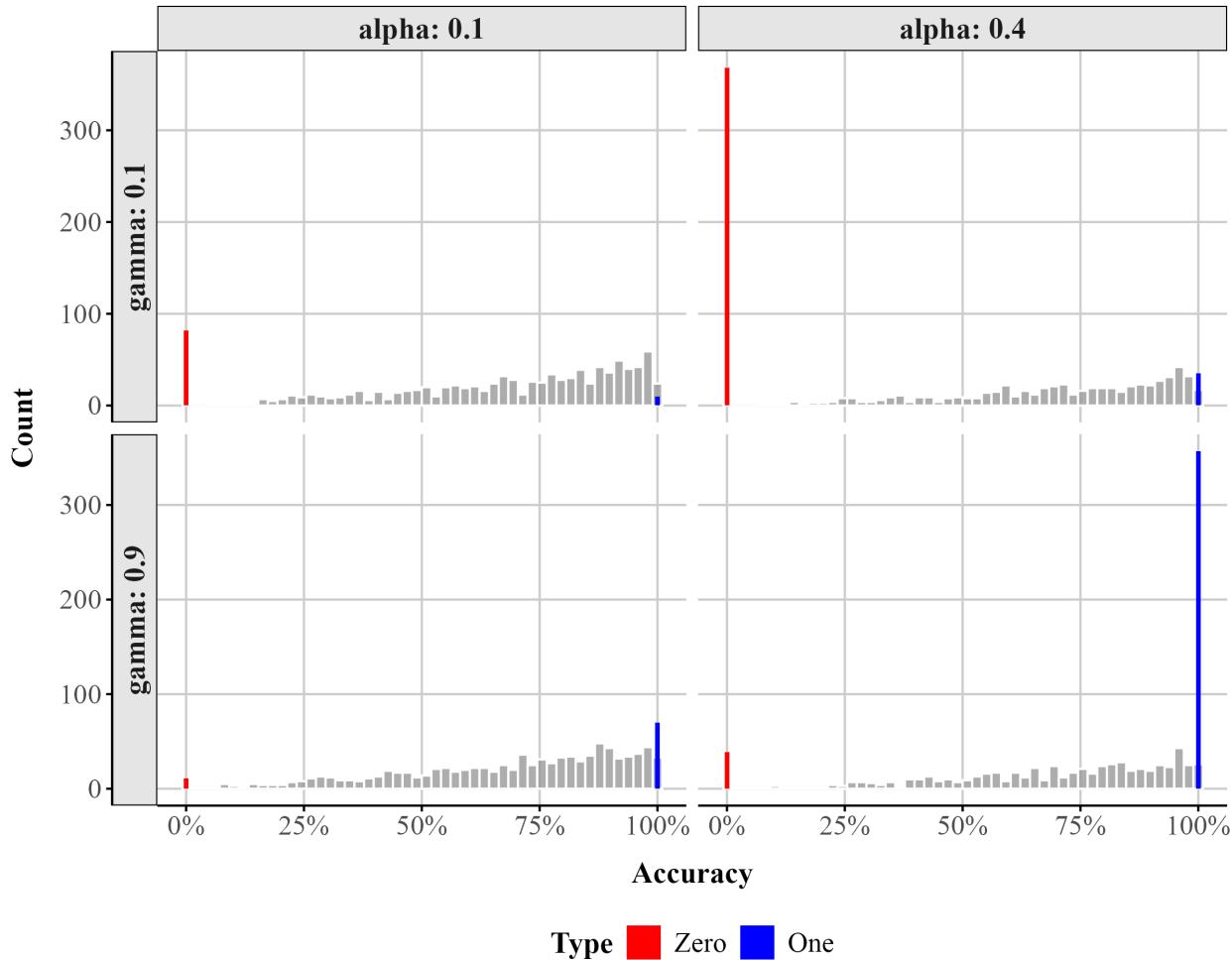
505 With `{marginaleffects}` we can choose `marginalize` over all the sub-models, averaged across the 0s,
 506 continuous responses, and 1s in the data, or we can model the parameters separately using the `dpar` argument
 507 like we did above setting it to: μ , ϕ , α , γ (see below). Using `avg_predictions()` and not setting `dpar` we
 508 can get the predicted probabilities across all the sub-models. We can also plot the overall difference between
 509 fluency and disfluency for the whole model with `plot_predictions()`.

510 In addition, we show below how one can extract the predicted probabilities and marginal effects for
 511 γ (and a similar process for any other model component, `zoi`, etc.):

⁸In cases where your data include exact 1s but no 0s, you can fit a one-inflated beta regression model in `{brms}` by setting the `coi` parameter to 1. This tells the model that all point masses occur at 1, rather than being split between 0 and 1. In other words, `coi = 1` assumes that any inflation in the data is due entirely to values at 1. In our data, we have exactly one value equal to 1^[^6]. While probably not significant to alter our findings, we can model 1s with a special type of model called the zero-one-inflated beta (ZOIB) model (Liu & Kong, 2015) if we believe that both 0s and 1s are distinct outcomes.

Figure 9

Simulated data from a ZOIB model illustrating the effects of the zero-one inflation parameter (α) and the conditional one-inflation parameter (γ).



512 Ordered Beta Regression

513 Looking at the output from the ZOIB model (Table 8), we can see how running a model like this
 514 can become fairly complex as it is fitting distinct sub-models for each component of the scale. The ability
 515 to consider 0s and 1s as distinct processes from continuous values comes at a price in terms of complexity
 516 and interpretability. A simplified version of the zero-one-inflated beta (ZOIB) model, known as ordered
 517 beta regression (Kubinec, 2022; see also Makowski et al., 2025 for a reparameterized version called the
 518 *beta-Gate* model), has been recently proposed. The ordered beta regression model exploits the fact that,
 519 for most analyses, the continuous values (between 0-1) and the discrete outcomes (e.g., 0 or 1) are *ordered*.
 520 For example, as a covariate x increases or decreases, we should expect the bounded outcome y to increase
 521 or decrease monotonically as well from 0 to (0, 1) to 1. The ZOIB model does not impose this restriction;
 522 a covariate could increase and the response y could increase in its continuous values while *simultaneously*
 523 decreasing at *both* end points.⁹ This complexity is not immediately obvious when fitting the ZOIB, nor is
 524 it a potential relationship that many scholars want to consider when examining how covariates influence a

⁹For a more complete description of this issue, we refer the reader to Kubinec (2022).

Listing 16 Fitting a ZOIB model with `brm()`.

```
# fit the zoib model
zoib_model <- bf(
  Accuracy ~ Fluency, # The mean of the 0-1 values, or mu
  phi ~ Fluency, # The precision of the 0-1 values, or phi
  zoi ~ Fluency, # The zero-or-one-inflated part, or alpha
  coi ~ Fluency, # The one-inflated part, conditional on the 1s, or gamma
  family = zero_one_inflated_beta()
)

fit_zoib <- brm(
  formula = zoib_model,
  data = fluency_data,
  file = here::here("manuscript", "models", "bayes_zoib_model")
)
```

Listing 17 Extracting predicted probabilities and marginal effects for conditional-one parameter

```
# get average predictions for coi param
coi_probs <- avg_predictions(fit_zoib, variables = "Fluency", dpar = "coi")
# get difference between the two conditions
coi_me <- avg_comparisons(fit_zoib, variables = c("Fluency"), dpar = "coi")
```

525 bounded scale.

526 To make the response ordered, the ordered beta regression model estimates a weighted combination
 527 of a standard beta regression model for continuous responses and a logit model for the discrete values of
 528 the response. By doing so, the amount of distinctiveness between the continuous responses and the discrete
 529 end points is a function of the data (and any informative priors) rather than strictly defined as fully distinct
 530 processes as in the ZOIB. For some datasets, the continuous and discrete responses will be fairly distinct,
 531 and in others less so.

532 The weights that average together the two parts of the outcome (i.e., discrete and continuous) are
 533 determined by cutpoints that are estimated in conjunction with the data in a similar manner to what is known
 534 as an ordered logit model. An in-depth explanation of ordinal regression is beyond the scope of this tutorial
 535 (Bürkner & Vuorre, 2019; but see Fullerton & Anderson, 2021). At a basic level, ordinal regression models
 536 are useful for outcome variables that are categorical in nature and have some inherent ordering (e.g., Likert
 537 scale items). To preserve this ordering, ordinal models rely on the cumulative probability distribution.
 538 Within an ordinal regression model it is assumed that there is a continuous but unobserved latent variable
 539 that determines which of k ordinal responses will be selected. For example on a typical Likert scale from
 540 'Strongly Disagree' to 'Strongly Agree', you could assume that there is a continuous, unobserved variable
 541 called 'Agreement'.

542 While we cannot measure Agreement directly, the ordinal response gives us some indication about
 543 where participants are on the continuous Agreement scale. $k - 1$ cutoffs are then estimated to indicate the
 544 point on the continuous Agreement scale at which your Agreement level is high enough to push you into the
 545 next ordinal category (say Agree to Strongly Agree). Coefficients in the model estimate how much differ-
 546 ent predictors change the estimated *continuous* scale (here, Agreement). Since there's only one underlying

547 process, there's only one set of coefficients to work with (proportional odds assumption).

548 In an ordered beta regression, three ordered categories are modeled: (1) exactly zero, (2) somewhere
 549 between zero and one, and (3) exactly one. In an ordered beta regression, (1) and (2) are modeled with
 550 cumulative logits, where one cutpoint is the the boundary between Exactly 0 and Between 0 and 1 and the
 551 other cutpoint is the boundary between *Between 0 and 1* and *Exactly 1*. The continuous values in the middle,
 552 0 to 1 (3), are modeled as a vanilla beta regression with parameters reflecting the mean response on the
 553 logit scale as we have described previously. Ultimately, employing cutpoints allows for a smooth transition
 554 between the bounds and the continuous values, permitting both to be considered together rather than modeled
 555 separately as the ZOIB requires.

556 The ordered beta regression model has shown to be more efficient and less biased than some of the
 557 methods discussed (Kubinec, 2022) herein and has seen increasing use across the biomedical and social
 558 sciences (Martin et al., 2024; Nouvian et al., 2023; Shrestha et al., 2024; Smith et al., 2024; Wilkes et al.,
 559 2024) because it produces only a single set of coefficient estimates in a similar manner to a standard beta
 560 regression or OLS.¹⁰

561 *Fitting an Ordered Beta Regression*

562 To fit an ordered beta regression in a Bayesian context we use the `{ordbetareg}` (Kubinec, 2023) pack-
 563 age. `{ordbetareg}` is a front-end to the `brms` package that we described earlier; in addition to the functions
 564 available in the package, most `brms` functions and plots, including the diverse array of regression model-
 565 ing options, will work with `{ordbetareg}` models. (We note that the `ordbeta` model is also available as a
 566 maximum-likelihood variant in the R package `{glmmTMB}`.) We first load the `{ordbetareg}` package (see
 567 Listing 18).

Listing 18 Load `{ordbetareg}`

```
library(ordbetareg)
```

568 The `{ordbetareg}` package uses `brms` on the front-end so all the arguments we used previously apply
 569 here. Instead of the `brm()` function we use `ordbetareg()`. To fit a model where dispersion does not vary
 570 as a function of fluency we can use the below code (see Listing 19).

Listing 19 Fitting ordered beta model with `ordbetareg()`

```
ord_fit_brms <- ordbetareg(  

  Accuracy ~ Fluency,  

  data = fluency_data,  

  file = here::here("manuscript","models", "bayes_ordbeta_model")  

)
```

571 However, if we want dispersion to vary as a function of fluency we can easily do that (see Listing 20).
 572 Note the addition of the `phi_reg` argument in `m.phi`. This argument allows us to include a model that
 573 explicitly models the dispersion parameter. Because we are modeling ϕ as a function of fluency, we set the
 574 the argument to `both`.

575 **Marginal Effects.** Table 8 presents the posterior summary. We can use `{marginaleffects}` to calcu-
 576 late differences on the response scale that average over (or marginalize over) all our parameters.

¹⁰Please note that there are other models available that can model this continuous process like the beta-gate model (Makowski et

Listing 20 Fitting ordered beta model with dispersion using `ordbetareg()`

```
ord_beta_phi <- bf(Accuracy ~ Fluency, phi ~ Fluency)

m.phi <- ordbetareg(
  ord_beta_phi,
  data = fluency_data,
  phi_reg = 'both',
  file = here::here("manuscript", "models", "bayes_ordbeta_phi_model")
)
```

Table 6*Marginal effect of fluency for ordered beta model*

Contrast	Mean	95% Cr.I	pd
Fluent - Disfluent	0.061	[-0.015, 0.138]	0.942

577 In Table 6 the credible interval is close enough to zero relative to its uncertainty that we can conclude
 578 there likely aren't differences between the conditions after taking dispersion and the 0s and 1s in our data
 579 into account.

580 **Cutpoints.** The model cutpoints are not reported by default in the summary output, but we can
 581 access them with the R package `posterior` (Bürkner et al., 2025) and the functions `as_draws` and
 582 `summary_draws`.

583 In Table 7, `cutzero` is the first cutpoint (the difference between 0 and continuous values) and `cutone`
 584 is the second cutpoint (the difference between the continuous values and 1). These cutpoints are on the
 585 logit scale and as such the numbers do not have a simple substantive meaning. In general, as the cutpoints
 586 increase in absolute value (away from zero), then the discrete/boundary observations are more distinct from
 587 the continuous values. This will happen if there is a clear gap or bunching in the outcome around the bounds.
 588 This type of empirical feature of the distribution may be useful to scholars if they want to study differences
 589 in how people perceive the ends of the scale versus the middle. It is possible, though beyond the scope of
 590 this article, to model the location of the cutpoints with hierarchical (non-linear) covariates in `brms`. In the
 591 most recent version of `ordbeta`, it is possible to test the influence of different factors on these boundaries.

592 Model Fit

593 The best way to visualize model fit is to plot the full predictive distribution relative to the original
 594 outcome. Because ordered beta regression is a mixed discrete/continuous model, a separate plotting function,

al., 2025) and the censored extended beta regression model (Kosmidis & Zeileis, 2025).

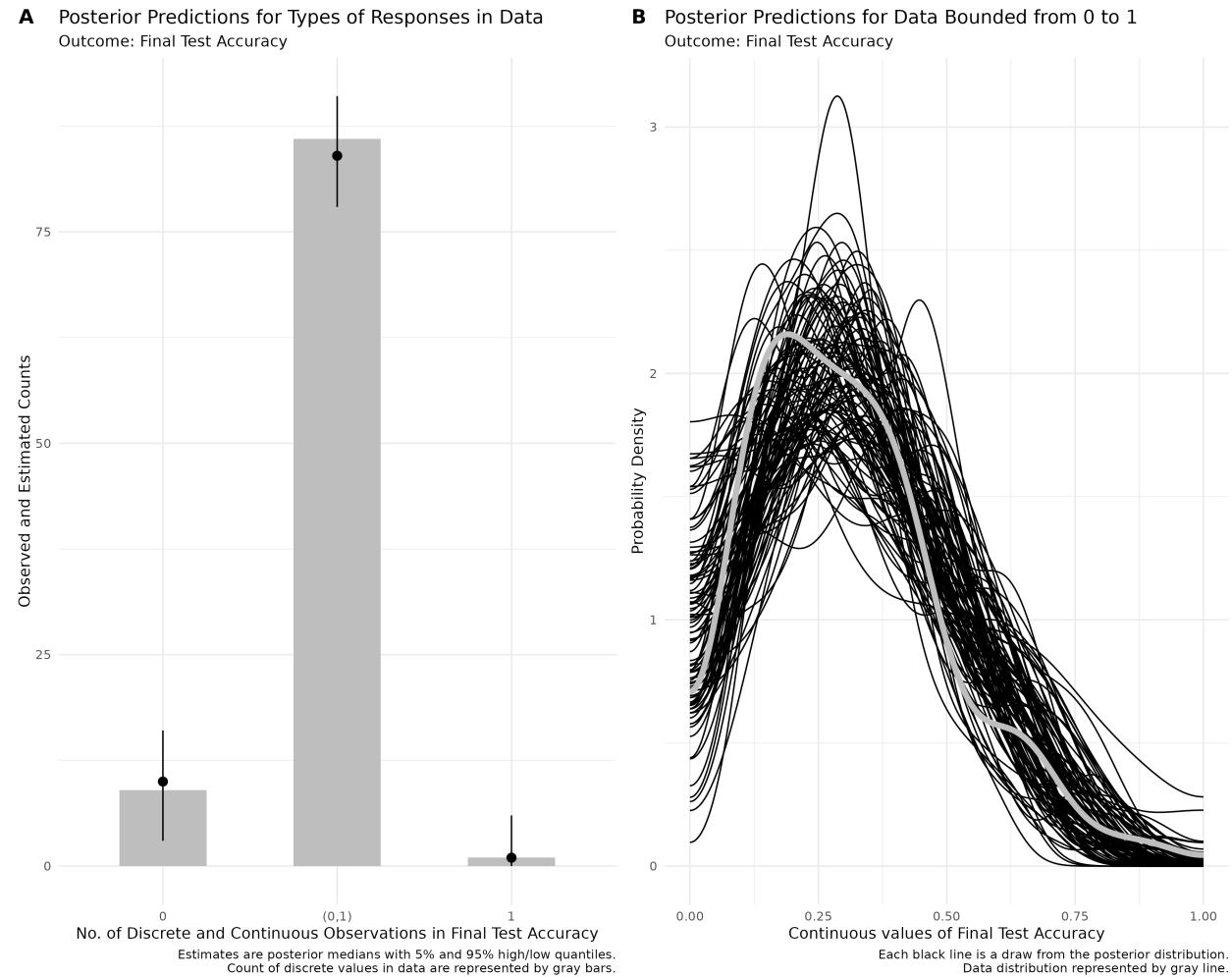
Table 7*Cutzero and cutone parameter summary*

Parameter	Mean	95% Cr.I
cutzero	-2.96	[-3.57, -2.41]
cutone	1.85	[1.64, 2.07]

595 `pp_check_ordbetareg`, is included in the `{ordbetareg}` package that accurately handles the unique features
 596 of this distribution. The default plot in `brms` will collapse these two features of the outcome together, which
 597 will make the fit look worse than it actually is. The `{ordbetareg}` function returns a list with two plots,
 598 `discrete` and `continuous`, which can either be printed and plotted or further modified as `{ggplot2}` objects
 599 (see Figure 10).

Figure 10

Posterior predictive check for ordered beta regression model. A. Discrete posterior check. B. Continuous posterior check.



600 The discrete plot, which is a bar graph, shows that the posterior distribution accurately captures the
 601 number of different types of responses (discrete or continuous) in the data. For the continuous plot shown as
 602 a density plot with one line per posterior draw, the model does a very good job at capturing the distribution.
 603 Overall, it is clear from the posterior distribution plot that the ordered beta model fits the data well.
 604 To fully understand model fit, both of these plots need to be inspected as they are conceptually distinct.

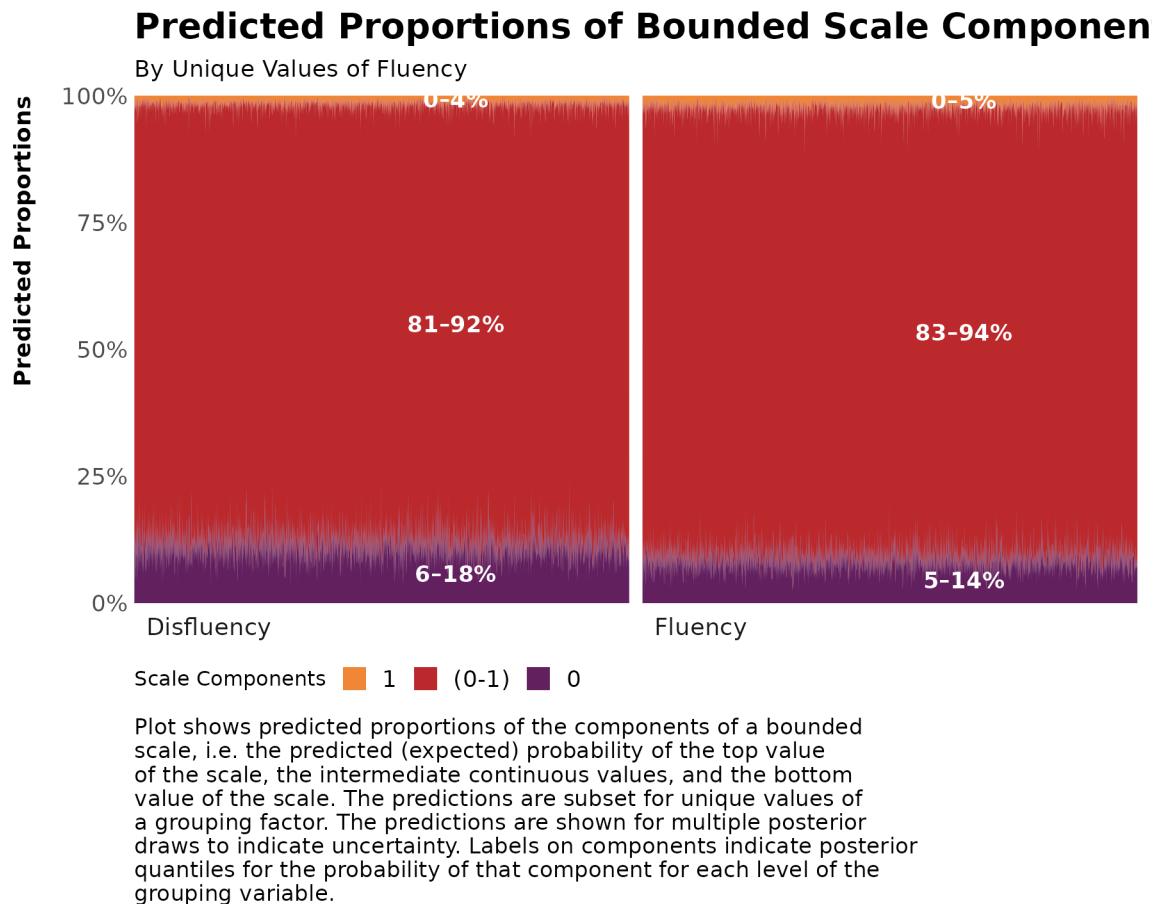
605 Model Visualization

606 `{ordbetareg}` provides a useful visualization function called `plot_heiss()` (Ye & Heiss, 2023) that
 607 can represent dispersion in the entire outcome as a function of discrete covariates. This function produces a

608 plot of predicted proportions across the range of our Fluency factor. In Figure 11 we get predicted proportions
 609 for Fluency across the bounded scale. Looking at the figure we can see there is much overlap between
 610 instructors in the middle portion (μ) . However, we do see some small differences at the zero bounds.

Figure 11

Heiss plot of predicted probabilities across the scale (0-100)



611 *Ordered Beta Scale*

612 In the `{ordbetareg}` function there is a `true_bound` argument. In cases where your data is not
 613 bounded between 0-1, this argument can be used to specify the bounds of the argument to fit the ordered
 614 beta regression. For example, the response data might be bounded between 1 and 7. If so, `{ordbetareg}` can
 615 model it within the [0,1] interval and `{ordbetareg}` will convert the model predictions back to the true bounds
 616 after estimation.

617

Discussion

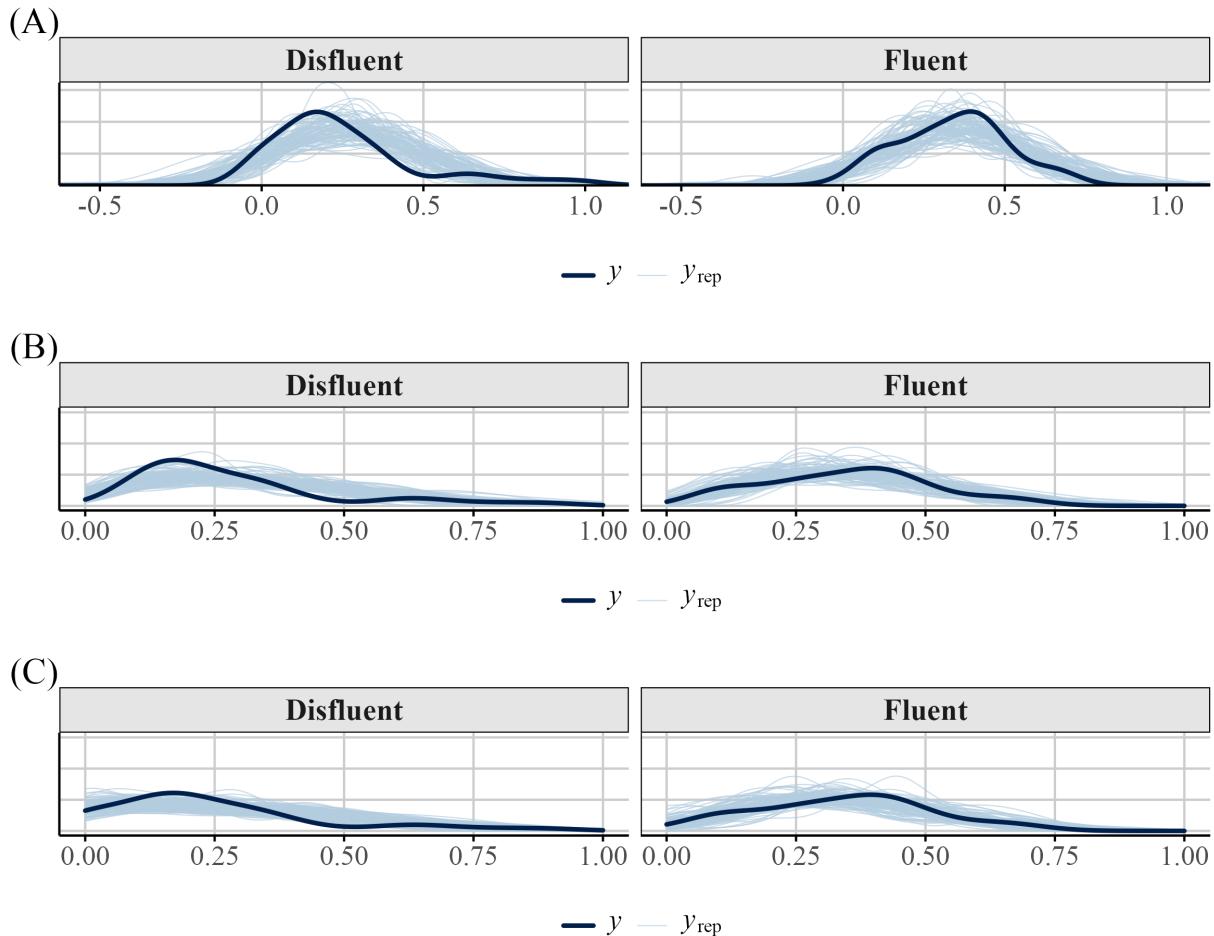
618 The use of beta regression in psychology, and the social sciences in general, is rare. With this tutorial,
 619 we hope to turn the tides. Beta regression models are an attractive alternative to models that impose unrea-
 620 listic assumptions like normality, linearity, homoscedasticity, and unbounded data. Beyond these models,
 621 there are a diverse array of different models that can be used depending on your outcome of interest.

Table 8*Bayesian regression summaries for each model*

Parameter	Stat	Bayesian LM	Beta Regression	ZIB	ZOIB	Ordered Beta
b_Intercept	Mean	0.257	-0.827	-0.830	-0.831	-0.865
	Cr.I	[0.202, 0.314]	[-1.089, -0.548]	[-1.096, -0.557]	[-1.1, -0.566]	[-1.125, -0.614]
	pd	1.000***	1.000***	1.000***	1.000***	1.000***
b_Fluency	Mean	0.085	0.198	0.201	0.204	0.261
	Cr.I	[-0.003, 0.169]	[-0.157, 0.539]	[-0.153, 0.547]	[-0.143, 0.55]	[-0.064, 0.592]
	pd	0.971	0.865	0.875	0.878	0.942
sigma	Mean	0.208	-	-	-	-
	Cr.I	[0.181, 0.24]	-	-	-	-
	pd	1.000***	-	-	-	-
b_phi_Intercept	Mean	-	1.601	1.601	1.599	1.616
	Cr.I	-	[1.184, 1.988]	[1.174, 1.994]	[1.187, 1.983]	[1.205, 1.991]
	pd	-	1.000***	1.000***	1.000***	1.000***
b_phi_Fluency	Mean	-	0.431	0.427	0.427	0.398
	Cr.I	-	[-0.142, 1]	[-0.167, 1.026]	[-0.164, 1.008]	[-0.196, 0.972]
	pd	-	0.934	0.922	0.927	0.913
b_zi_Intercept	Mean	-	-	-1.652	-	-
	Cr.I	-	-	[-2.45, -0.94]	-	-
	pd	-	-	1.000***	-	-
b_zi_Fluency	Mean	-	-	-2.100	-	-
	Cr.I	-	-	[-4.444, -0.368]	-	-
	pd	-	-	0.991**	-	-
b_zoi_Intercept	Mean	-	-	-	-1.545	-
	Cr.I	-	-	-	[-2.297, -0.864]	-
	pd	-	-	-	1.000***	-
b_zoi_Fluency	Mean	-	-	-	-2.258	-
	Cr.I	-	-	-	[-4.716, -0.455]	-
	pd	-	-	-	0.995***	-
b_coi_Intercept	Mean	-	-	-	-2.084	-
	Cr.I	-	-	-	[-4.526, -0.36]	-
	pd	-	-	-	0.993**	-

Figure 12

The plots show 100 posterior predicted distributions with the label y_{rep} (light blue), the distribution of accuracy as function of fluency in dark blue for regular regression (A), beta regression (B), ZIB (C) and ZOIB (D).



Throughout this tutorial our main aim was to help guide researchers in running analyses with proportional or percentage outcomes using beta regression and some of its alternatives. In the current example, we used real data from Wilford et al. (2020) and discussed how to fit these models in R, interpret model parameters, extract predicted probabilities and marginal effects, and visualize the results using the `{marginaleffects}` package.

Comparing our analysis with that of Wilford et al. (2020), we demonstrated that using a traditional approach (e.g., t -test) to analyze mean accuracy data can lead to biased inferences. Although we successfully reproduced one of their key findings, our use of beta regression and its extensions revealed important nuances in the results. With a traditional beta regression model—which accounts for both the mean and the precision (dispersion)—we observed similar effects of instructor fluency on performance. However, the standard beta model does not accommodate boundary values (i.e., 0s and 1s).

When we applied a ZIB model, which explicitly accounts for structural 0s, we found no effect of fluency on the mean (μ) part of the model. Instead, the effect of fluency emerged in the structural zero (inflated zero; α) component. This pattern was consistent when using a zero-one-inflated beta (ZOIB) model.

636 Furthermore, we fit an ordered beta regression model (Kubinec, 2022), which appropriately models the full
637 range of values, including 0s and 1s. Here, we did not observe a reliable effect of fluency on the mean once
638 we accounted for dispersion.

639 These analyses emphasize the importance of fitting a model that aligns with the nature of the data.
640 The simplest and recommended approach when dealing with data that contains 0s and/or 1s is to fit an ordered
641 beta model, assuming the process is truly continuous. However, if you believe the process is distinct in nature,
642 a ZIB or ZOIB model might be a better choice. Ultimately, this decision should be guided by theory.

643 For instance, if we believe fluency influences the boundaries (0 and 1), we might want to model
644 this process separately using a ZIB or ZOIB. With the current dataset, fluency might affect specific aspects
645 of performance (such as the likelihood of complete failure) rather than general performance levels. This
646 effect could be due to participant disengagement during the disfluent lecture. If students fail to pay attention
647 because of features of disfluency, they may miss relevant information, leading to a floor effect at the test.
648 Following from this, disfluency would be expected to influence the boundary (0) and not the continuous part
649 of the model. If this is the case, we would want to model this appropriately. However, if we believe fluency
650 effects general performance levels (the continuous part), a model that takes in to account the entire process
651 accounting for the 0s and 1s might be appropriate.

652 In the discussion section of Wilford et al. (2020), they were unable to offer a tenable explanation for
653 performance differences based on instructor fluency. A model that accounts for the excess 0s in the dataset
654 provides one testable explanation: watching a disfluent lecture may lead to lapses in attention, resulting
655 in poorer performance in that group. These lapses, in turn, contribute to the observed differences in the
656 fluent condition. This modeling approach opens a promising avenue for future research—one that would have
657 remained inaccessible otherwise.

658 Not everyone will be eager to implement the techniques discussed herein. In such cases, the key
659 question becomes: What is the least problematic approach to handling proportional data? One reasonable
660 option is to fit multiple models tailored to the specific characteristics of your data. For example, if your data
661 contain 0s, you might fit two models: a traditional linear model excluding the 0s, and a logistic model to
662 account for the zero versus non-zero distinction. If your data contain both 0s and 1s, you could fit separate
663 models for the 0s and 1s in addition to the OLS model. There are many defensible strategies to choose from
664 depending on the context. However, we do not recommend transforming the values of your data (e.g., 0s to
665 .01 and 1s to .99) or ignoring the properties of your data simply to fit traditional statistical models.

666 In this tutorial, we demonstrated how to analyze these models from a Bayesian perspective. While we
667 recognize that not everyone identifies as a Bayesian, implementing these models using a Bayesian framework
668 is relatively straightforward—it requires only a single package, lowering the barrier to entry. For those who
669 prefer frequentist analyses, several R packages are available. For example, the `{betareg}` package (Cribari-
670 Neto & Zeileis, 2010) `{glmmTMB}` (Brooks et al., 2017) and `{gamlss}` (2005) are nice options. To this end,
671 I have included supplemental materials demonstrating how to use frequentist packages to analyze the data
672 presented herein.

673 Conclusion

674 Overall, this tutorial emphasizes the importance of modeling the data you have. Although the ex-
675 ample provided is relatively simple (a one-factor model with two levels), we hope it demonstrates that even
676 with a basic dataset, there is much nuance in interpretation and inference. Properly modeling your data
677 can lead to deeper insights, far beyond what traditional measures might offer. With the tools introduced in
678 this tutorial, researchers now have the means to analyze their data effectively, uncover patterns, make ac-
679 curate predictions, and support their findings with robust statistical evidence. By applying these modeling
680 techniques, researchers can improve the validity and reliability of their studies, ultimately leading to more
681 informed decisions and advancements in their respective fields.

References

- 683 Arel-Bundock, V. (2024). *Marginaleffects: Predictions, comparisons, slopes, marginal means, and hypothesis tests*. <https://CRAN.R-project.org/package=marginaleffects>

684

685 Arel-Bundock, V., Greifer, N., & Heiss, A. (2024). How to interpret statistical models using marginaleffects
686 for R and Python. *Journal of Statistical Software*, 111(9), 1–32. <https://doi.org/10.18637/jss.v111.i09>

687 Bartlett, M. S. (1936). The square root transformation in analysis of variance. *Journal of the Royal Statistical
688 Society Series B: Statistical Methodology*, 3(1), 68–78. <https://doi.org/10.2307/2983678>

689 Bendixen, T., & Purzycki, B. G. (2023). Cognitive and cultural models in psychological science: A tutorial
690 on modeling free-list data as a dependent variable in bayesian regression. *Psychological Methods*. <https://doi.org/10.1037/met0000553>

691 Brooks, M. E., Kristensen, K., van, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Maechler,
692 M., & Bolker, B. M. (2017). *{glmmTMB} balances speed and flexibility among packages for zero-inflated
693 generalized linear mixed modeling*. 9. <https://doi.org/10.32614/RJ-2017-066>

694 Bürkner, P.-C. (2017). *{Brms}: An {r} package for {bayesian} multilevel models using {stan}*. 80. <https://doi.org/10.18637/jss.v080.i01>

695 Bürkner, P.-C., Gabry, J., Kay, M., & Vehtari, A. (2025). *posterior: Tools for working with posterior distri-
696 butions*. <https://mc-stan.org/posterior/>

697 Bürkner, P.-C., & Vuorre, M. (2019). Ordinal regression models in psychology: A tutorial. *Ad-
698 vances in Methods and Practices in Psychological Science*, 2(1), 77–101. <https://doi.org/10.1177/2515245918823199>

699 Carpenter, S. K., Wilford, M. M., Kornell, N., & Mullaney, K. M. (2013). Appearances can be deceiving:
700 Instructor fluency increases perceptions of learning without increasing actual learning. *Psychonomic
701 Bulletin & Review*, 20(6), 1350–1356. <https://doi.org/10.3758/s13423-013-0442-z>

702 Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*, rev. ed. Lawrence Erlbaum Asso-
703 ciates, Inc.

704 Coretta, S., & Bürkner, P.-C. (2025). *Bayesian beta regressions with brms in r: A tutorial for phoneticians*.
705 https://doi.org/10.31219/osf.io/f9rqp_v1.

706 Costello, T. H. et al. (2024). Durably reducing conspiracy beliefs through dialogues with AI. *Science*, 385,
707 eadq1814. <https://doi.org/10.1126/science.adq1814>

708 Cribari-Neto, F., & Zeileis, A. (2010). *Beta regression in {r}*. 34. <https://doi.org/10.18637/jss.v034.i02>

709 Dolstra, E., & contributors, T. N. (2006). *Nix* [Computer software]. <https://nixos.org/>

710 Ferrari, S., & Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of
711 Applied Statistics*, 31(7), 799–815. <https://doi.org/10.1080/0266476042000214501>

712 Fullerton, A. S., & Anderson, K. F. (2021). Ordered regression models: A tutorial. *Prevention Science*,
713 24(3), 431–443. <https://doi.org/10.1007/s11121-021-01302-y>

714 Gabry, J., Češnovar, R., Johnson, A., & Brodner, S. (2024). *Cmdstanr: R interface to 'CmdStan'*. <https://mc-stan.org/cmdstanr/>

715 Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data
716 analysis* (Third). CRC. <https://stat.columbia.edu/~gelman/book/>

717 Heiss, A. (2021). *A guide to modeling proportions with bayesian beta and zero-inflated beta regression
718 models*. <http://dx.doi.org/10.59350/7p1a4-0tw75>

719 James, A. N., Ryskin, R., Hartshorne, J. K., Backs, H., Bala, N., Barcenas-Meade, L., Bhattacharai, S., Charles,
720 T., Copoulos, G., Coss, C., Eisert, A., Furuhashi, E., Ginell, K., Guttman-McCabe, A., Harrison, E.
721 (Chaz.), Hoban, L., Hwang, W. A., Iannetta, C., Koenig, K. M., ... Leeuw, J. R. de. (2025). What
722 paradigms can webcam eye-tracking be used for? Attempted replications of five cognitive science exper-
723 iments. *Collabra: Psychology*, 11(1). <https://doi.org/10.1525/collabra.140755>

724 Johnson, A., Ott, M., & Dogucu, M. (2022). *Bayes rules!: An introduction to applied bayesian modeling*.

- 729 Routledge & CRC Press.
- 730 Kneib, T., Silbersdorff, A., & Säfken, B. (2023). Rage against the mean – a review of distributional regression
731 approaches. *Econometrics and Statistics*, 26, 99–123. <https://doi.org/10.1016/j.ecosta.2021.07.006>
- 732 Kong, E. J., & Edwards, J. (2016). Individual differences in categorical perception of speech: Cue weighting
733 and executive function. *Journal of Phonetics*, 59, 40–57. <https://doi.org/10.1016/j.wocn.2016.08.006>
- 734 Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories. *Psychological Science*, 19(6), 585–
735 592. <https://doi.org/10.1111/j.1467-9280.2008.02127.x>
- 736 Kosmidis, I., & Zeileis, A. (2025). Extended-support beta regression for [0, 1] responses. *Journal of the
737 Royal Statistical Society Series C: Applied Statistics*. <https://doi.org/10.1093/rsscc/qlaf039>
- 738 Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, 142(2), 573–603. <https://doi.org/10.1037/a0029146>
- 740 Kruschke, J. K. (2015). *Doing bayesian data analysis: A tutorial with r, JAGS, and stan* (2nd ed.). Academic
741 Press.
- 742 Kruschke, J. K. (2018). Rejecting or accepting parameter values in bayesian estimation. *Advances in Methods
743 and Practices in Psychological Science*, 1(2), 270–280. <https://doi.org/10.1177/2515245918771304>
- 744 Kubinec, R. (2022). Ordered beta regression: A parsimonious, well-fitting model for continuous data with
745 lower and upper bounds. *Political Analysis*, 31(4), 519–536. <https://doi.org/10.1017/pan.2022.20>
- 746 Kubinec, R. (2023). *Ordbetareg: Ordered beta regression models with 'brms'*. <https://CRAN.R-project.org/package=ordbetareg>
- 748 Lenth, R. V. (2025). *Emmeans: Estimated marginal means, aka least-squares means*. <https://doi.org/10.32614/CRAN.package.emmeans>
- 750 Liu, F., & Kong, Y. (2015). Zoib: An r package for bayesian inference for beta regression and zero/one
751 inflated beta regression. *The R Journal*, 7(2), 34. <https://doi.org/10.32614/rj-2015-019>
- 752 Lüdecke, D., Ben-Shachar, M. S., Patil, I., Wiernik, B. M., Bacher, E., Thériault, R., & Makowski,
753 D. (2022). *Easystats: Framework for easy statistical modeling, visualization, and reporting*. <https://easystats.github.io/easystats/>
- 755 Makowski, D., Ben-Shachar, M. S., Chen, S. H. A., & Lüdecke, D. (2019). Indices of effect existence and
756 significance in the bayesian framework. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.02767>
- 758 Makowski, D., Ben-Shachar, M., & Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty,
759 existence and significance within the bayesian framework. *Journal of Open Source Software*, 4(40), 1541.
760 <https://doi.org/10.21105/joss.01541>
- 761 Makowski, D., Neves, A., & Field, A. P. (2025). *Introducing the choice-confidence (CHOCO) model for
762 bimodal data from subjective ratings: Application to the effect of attractiveness on reality beliefs about
763 AI-generated faces*. https://doi.org/10.31234/osf.io/z68v3_v1
- 764 Marsman, M., & Wagenmakers, E.-J. (2016). Three Insights from a Bayesian Interpretation of the One-Sided
765 P Value. *Educational and Psychological Measurement*. <https://doi.org/10.1177/0013164416669201>
- 766 Martin, K., Cornero, F. M., Clayton, N. S., Adam, O., Obin, N., & Dufour, V. (2024). Vocal complexity in a
767 socially complex corvid: Gradation, diversity and lack of common call repertoire in male rooks. *Royal
768 Society Open Science*, 11(1), 231713. <https://doi.org/10.1098/rsos.231713>
- 769 McCabe, C. J., Halvorson, M. A., King, K. M., Cao, X., & Kim, D. S. (2021). Interpreting interaction effects
770 in generalized linear models of nonlinear probabilities and counts. *Multivariate Behavioral Research*,
771 57(2-3), 243–263. <https://doi.org/10.1080/00273171.2020.1868966>
- 772 McElreath, R. (2020). *Statistical rethinking: A bayesian course with examples in r and STAN* (2nd ed.).
773 Chapman; Hall/CRC. <https://doi.org/10.1201/9780429029608>
- 774 Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical
775 Society. Series A (General)*, 135(3), 370–384. <https://doi.org/10.2307/2344614>
- 776 Nouvian, M., Foster, J. J., & Weidenmüller, A. (2023). Glyphosate impairs aversive learning in bumble-

- 777 bees. *Science of The Total Environment*, 898, 165527. <https://www.sciencedirect.com/science/article/pii/S0048969723041505>
- 778 Paolino, P. (2001). Maximum likelihood estimation of models with beta-distributed dependent variables. *Political Analysis*, 9(4), 325–346. <https://doi.org/10.1093/oxfordjournals.pan.a004873>
- 781 Pfadt, J. M., Bartoš, F., Godmann, H. R., Waaijers, M., Groot, L., Heo, I., & Wagenmakers, E. (2025).
782 A methodological metamorphosis: The rapid rise of bayesian inference and open science practices in
783 psychology. *Preprint*. https://doi.org/10.31234/osf.io/ck3js_v1
- 784 R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical
785 Computing. <https://www.R-project.org/>
- 786 R. A. Rigby, & D. M. Stasinopoulos. (2005). *Generalized additive models for location, scale and shape,(with
787 discussion)*. 54, 507–554.
- 788 Rodrigues, B., & Baumann, P. (2025). *Rix: Reproducible data science environments with 'nix'*. <https://docs.ropensci.org/rix/>
- 790 Rohrer, J. M., & Arel-Bundock, V. (2025). *Models as prediction machines: How to convert confusing coef-
791 ficients into clear quantities*. https://doi.org/10.31234/osf.io/g4s2a_v1
- 792 Shrestha, S., Sigdel, K., Pokharel, M., & Columbus, S. (2024). Big five traits predict between- and within-
793 person variation in loneliness. *European Journal of Personality*, 08902070241239834. <https://doi.org/10.1177/08902070241239834>
- 795 Sladekova, M., & Field, A. P. (2024). *In search of unicorns: Assessing statistical assumptions in real psy-
796 chology datasets*. <https://doi.org/10.31234/osf.io/4rznt>
- 797 Smith, K. E., Panlilio, L. V., Feldman, J. D., Grundmann, O., Dunn, K. E., McCurdy, C. R., Garcia-
798 Romeu, A., & Epstein, D. H. (2024). Ecological momentary assessment of self-reported kratom use,
799 effects, and motivations among US adults. *JAMA Network Open*, 7(1), e2353401. <https://doi.org/10.1001/jamanetworkopen.2023.53401>
- 801 Smithson, M., & Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with
802 beta-distributed dependent variables. *Psychological Methods*, 11(1), 54–71. <https://doi.org/10.1037/1082-989x.11.1.54>
- 804 Team, S. D. (2023). *Stan: A probabilistic programming language*. <https://mc-stan.org>
- 805 Toftness, A. R., Carpenter, S. K., Geller, J., Lauber, S., Johnson, M., & Armstrong, P. I. (2017). Instructor
806 fluency leads to higher confidence in learning, but not better learning. *Metacognition and Learning*,
807 13(1), 1–14. <https://doi.org/10.1007/s11409-017-9175-0>
- 808 Vuorre, M. (2019, February 18). *How to analyze visual analog (slider) scale data?* <https://vuorre.com/posts/2019-02-18-analyze-analog-scale-ratings-with-zero-one-inflated-beta-models>
- 810 Wickham, H., Çetinkaya-Rundel, M., & Grolemund, G. (2023). *R for Data Science: Import, Tidy, Transform,
811 Visualize, and Model Data*. O'Reilly. <https://r4ds.hadley.nz/>
- 812 Wilford, M. M., Kurpad, N., Platt, M., & Weinstein-Jones, Y. (2020). Lecturer fluency can impact students'
813 judgments of learning and actual learning performance. *Applied Cognitive Psychology*, 34(6), 1444–
814 1456. <https://doi.org/10.1002/acp.3724>
- 815 Wilkes, L. N., Barner, A. K., Keyes, A. A., Morton, D., Byrnes, J. E. K., & Dee, L. E. (2024). Quantifying
816 co-extinctions and ecosystem service vulnerability in coastal ecosystems experiencing climate warming.
817 *Global Change Biology*, 30(7), e17422. <https://doi.org/10.1111/gcb.17422>
- 818 Witherby, A. E., & Carpenter, S. K. (2022). The impact of lecture fluency and technology fluency on students'
819 online learning and evaluations of instructors. *Journal of Applied Research in Memory and Cognition*,
820 11(4), 500–509. <https://doi.org/10.1037/mac0000003>
- 821 Ye, M., & Heiss, A. (2023). Enforcing Boundaries: China's Overseas NGO Law and Operational Con-
822 straints for Global Civil Society. *Working Paper*. <https://stats.andrewheiss.com/compassionate-clam/manuscript/output/manuscript.html>
- 823